

IO para alunos de LMAC

Duarte Maia

duarte.nascimento@tecnico.ulisboa.pt

Conteúdo

1	Introdução	2
1.1	Introdução à versão 1	2
1.2	Introdução à versão 2	3
2	Pré-requisitos	3
3	Problemas de Otimização Linear	3
3.1	Introdução	3
3.2	Tradução	4
3.2.1	Substituição de equivalências	5
3.2.2	Adição de positividades (Forma canónica)	6
3.2.3	Desigualdades para igualdades (Forma padrão)	8
3.2.4	Padrão para Canónico	9
4	Noções Topológicas	9
4.1	Intuição geométrica	9
4.2	Factos introdutórios	10
5	Álgebra Linear	12
5.1	Intuição geométrica	12
5.2	Geometria	13
5.3	Vértices de conjuntos admissíveis	15
5.4	O método dos vetores básicos	20
6	Cones	22
6.1	Introdução	22
6.2	Método das linhas ativas (Parte 1)	23
6.3	O Lema de Farkas (Parte 1)	26
6.4	Cones primitivos	27
6.5	O Lema de Farkas (Parte 2)	30
6.6	Método das linhas ativas (Parte 2)	32

7	Dualidade	33
7.1	Motivação	33
7.2	Introdução à dualidade	34
7.3	Algumas palavras sobre o dual	36
7.4	Propriedades básicas do dual	37
7.5	Outros duais	41
8	Simplexo	44
8.1	Introdução	44
8.2	Motivação	44
8.3	Concretização	47
8.4	Pontas soltas	49
A	Conceitos de Álgebra Linear	51
B	Conceitos Topológicos	51

1 Introdução

1.1 Introdução à versão 1

Uma das minhas coisas favoritas de fazer quando pego num livro novo de matemática é ler a introdução. Sei que há pessoal que passa isso à frente, porque não serve realmente para nada. No entanto, é algo que me agrada, porque sei que é a única parte do livro que vou ler a um ritmo de mais de 10 palavras por minuto.

Esta versão do “IO para alunos de LMAC” (que era originalmente chamado “IO Done Right” mas mudei o nome porque não queria insultar demasiado a Sernadas) é uma versão bastante primordial. Ainda não reli metade disto, pelo que inevitavelmente há imensas partes que estarão bastante insatisfatórias. Assim sendo, apelo ao leitor que mande para o meu e-mail (duarte.nascimento@tecnico.ulisboa.pt) quaisquer críticas que tenha, desde ‘Falta aqui um ponto final’ até ‘Este teorema é falso’, e tudo no meio, incluindo (e desejando-se) coisas para o efeito de ‘esta prova está mal explicada’. O objetivo é a compreensão, pelo que desejo fazer um texto acessível a qualquer aluno de LMAC que saiba ler.

Uma nota final: a secção de pré-requisitos está, neste momento, muito desarrumada. Estou eventualmente a planear melhorá-la, e escrever anexos a tratar dos conceitos básicos topológicos necessários. De qualquer das formas, se o leitor sentir que há algum conhecimento base que lhe falte, ou que algo está mal explicado, é favor assinalar e informar-me, para o problema ser retificado em versões futuras.

Sem mais atrasos, eis a sebenta em que estou a trabalhar desde o verão de 2018.

-Duarte Maia

1.2 Introdução à versão 2

To do.

2 Pré-requisitos

Vou aqui escrevendo coisas à medida que elas são necessárias.

Obviamente é preciso saber um pouco de cálculo I e AL.

É recomendado (i.e. necessário) que o leitor saiba os seguintes conceitos topológicos em \mathbb{R}^n :

Conceito de conjunto aberto e conjunto fechado; interior, fronteira e exterior de um conjunto; função contínua; teorema de Bolzano e Weierstrass em \mathbb{R}^n , definição topológica de continuidade ($f : D \rightarrow \mathbb{R}^m$ é contínua sse para todo conjunto aberto A se tem $f^{-1}(A)$ é aberto relativamente a D sse para todo conjunto fechado A se tem $f^{-1}(A)$ é fechado relativamente a D)

No que se segue vão ser usados alguns factos sobre estes conceitos que não vão ser justificados, visto que caem no âmbito da cadeira de Cálculo II. Estes serão devidamente assinalados, e o leitor é encorajado a prová-los se não estiver familiar com eles.

Referências para AL e CDI:

- Linear Algebra Done Right, Sheldon Axler
- Calculus on Manifolds, Michael Spivak

Referência na qual me baseei para o conteúdo de IO: as notas da cadeira de IO escritas por: Diogo Gomes, Amílcar Sernadas, Cristina Sernadas, João Rasga, Paulo Mateus; estas notas encontram-se na página do Fénix da cadeira de IO.

3 Problemas de Otimização Linear

3.1 Introdução

Um problema de otimização é um tipo específico de problema em que o objetivo é minimizar ou maximizar uma certa função (função-objetivo ou pontuação), dentro de um certo domínio, usualmente (no contexto desta cadeira) um subconjunto de \mathbb{R}^n parametrizado por um conjunto de (in)equações.

Mais concretamente, o tipo de problemas com que se lida nesta cadeira são problemas de otimização linear. Estes são problemas em que a função-objetivo é uma função linear $\mathbb{R}^n \rightarrow \mathbb{R}$ e todas as condições no domínio são da forma $ax \leq b$, $ax = b$ ou $ax \geq b$, onde a é um vetor-linha e b é um escalar.

Definição 1. Um *problema de otimização linear* (normalmente abreviado a *pol*) é um problema da forma:

Objetivo: maximizar/minimizar (em função de $x \in \mathbb{R}^n$) o valor de $c_1x_1 + c_2x_2 + \dots + c_nx_n$

Restrições: x tem de obedecer a todas as seguintes igualdades

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n = b_2 \\ \vdots \\ a_{p1}x_1 + a_{p2}x_2 + \dots + a_{pn}x_n = b_p \end{cases}$$

E todas as seguintes desigualdades

$$\begin{cases} a'_{11}x_1 + a'_{12}x_2 + \dots + a'_{1n}x_n \leq b'_1 \\ \vdots \\ a'_{q1}x_1 + a'_{q2}x_2 + \dots + a'_{qn}x_n \leq b'_q \\ \vdots \\ a''_{11}x_1 + a''_{12}x_2 + \dots + a''_{1n}x_n \geq b''_1 \\ \vdots \\ a''_{r1}x_1 + a''_{r2}x_2 + \dots + a''_{rn}x_n \geq b''_r \end{cases}$$

Infelizmente, isto é algo trabalhoso de escrever. Como tal, normalmente estas condições são escritas de forma mais compacta usando a linguagem da álgebra linear.

Notação. Dados dois vetores $x, y \in \mathbb{R}^n$, dizemos $x \leq y$ se $x_i \leq y_i$ para todo i .

Notação. Considere-se o pol escrito acima. Este é normalmente escrito da seguinte forma:

(Supõe tratar-se de um problema de maximização; caso contrário escreva-se min no lugar de max)

Seja c o vetor linha $[c_1 \ c_2 \ \dots \ c_n]$,

A a matriz $p \times n$ cujo i, j -ésimo elemento é a_{ij} , e considerações análogas para A' a A'' , com a'_{ij} e a''_{ij} respetivamente

b o vetor coluna (b_1, b_2, \dots, b_p) e análogamente para b' e b'' .

$$\begin{cases} \max_x cx \\ Ax = b \\ A'x \leq b' \\ A''x \geq b'' \end{cases}$$

Normalmente, não é útil considerar um tipo de problema tão geral, pelo que discutimos em baixo formas de passar de um tipo de problemas para outros.

3.2 Tradução

Considere-se os seguintes problemas de otimização linear:

$$\begin{cases} \max_x 2x_1 - 3x_2 \\ 5x_1 + x_2 = 6 \\ 2x_1 + 0x_2 \leq 2 \end{cases} \quad \begin{cases} \min_x -2x_1 + 3x_2 \\ 5x_1 + x_2 \leq 6 \\ 5x_1 + x_2 \geq 6 \\ -2x_1 + 0x_2 \geq -2 \end{cases}$$

Apesar de terem uma aparência diferente, alguma inspeção leva à conclusão que estes são, na realidade, exatamente o mesmo problema. Ou seja, o mesmo problema pode ser representado de mais do que uma forma diferente. Isto leva à possibilidade de considerar ‘formas canónicas’ de expressar os problemas, que sejam mais fáceis de estudar. Se, por exemplo, dissermos que todos os problemas da forma XYZ podem ser resolvidos fazendo ABC, e arranjarmos forma de traduzir qualquer pol para esta forma XYZ, estamos numa boa situação.

Notação. Dado um pol P , o conjunto de $x \in \mathbb{R}^n$ que satisfazem as suas condições é denominado de *conjunto admissível de P* , representado por X_P , ou só X se o pol em questão for óbvio de contexto.

Da mesma forma, o *conjunto solução de P* , S_P , ou S se P é claro de contexto, é o conjunto de $x \in X$ que maximizam, de facto, a função objetivo. Por outras palavras,

$$S_P = \{ x \in X_P \mid \forall_{y \in X_P} cx \geq cy \}$$

Isto no caso de maximização. No caso de minimização, a desigualdade deve estar trocada.

Há várias formas possíveis de transformar um problema noutra.

3.2.1 Substituição de equivalências

Por exemplo, a condição de igualdade $a = b$ pode ser expressa como duas desigualdades: $a \leq b$ e $a \geq b$. Assim sendo, sabemos à partida que qualquer pol que nos venha à cabeça pode ser expresso só com condições de desigualdade.

Para mais, a desigualdade $a \geq b$ é equivalente a $-a \leq -b$. Logo, podemos sempre assumir que as condições que restringem um pol são sempre da forma $Ax \leq b$.

Podemos também assumir sem perda de generalidade que um problema de otimização linear é um problema de maximização, pois minimizar cx é o mesmo que maximizar $-cx$.

Logo, dado um pol qualquer, podemos sempre assumir, sem perda de generalidade, que é da forma

$$\begin{cases} \max_x cx \\ Ax \leq b \end{cases}$$

A esta forma vamos-nos referir, no futuro, como ‘forma semicanónica’, na falta de melhor nome. Sugestões são aceites.

O que foi até agora descrito é o tipo mais simples possível de tradução de um problema a outro: trocar as condições por condições equivalentes, e substituir ‘maximizar f ’ por ‘minimizar $-f$ ’.

Este tipo de tradução tem a propriedade que não muda o conjunto admissível e conjunto solução, mas há traduções que não são tão triviais.

3.2.2 Adição de positivities (Forma canónica)

É, em muitos casos, útil estudar um problema de otimização em que sabemos que todas as variáveis são não-negativas. Assim sendo, vamos agora ver que podemos sempre, sem perda de generalidade, assumir que o nosso pol é da forma

$$\begin{cases} \max_x cx \\ Ax \leq b \\ x \geq 0 \end{cases}$$

O truque é modificar o conjunto em que estamos.

Dado um número real x , define-se a sua parte positiva, x^+ , como

$$x^+ = \begin{cases} x & \text{se } x \geq 0 \\ 0 & \text{caso contrário} \end{cases}$$

e a parte negativa, x^- , como

$$x^- = \begin{cases} 0 & \text{se } x \geq 0 \\ -x & \text{caso contrário} \end{cases}$$

Dado um vetor $x \in \mathbb{R}^n$, define-se x^+ como o vetor das partes positivas de x , e x^- como o vetor das partes negativas. É claro que $x = x^+ - x^-$, e $x^+, x^- \geq 0$. Isto permite-nos escrever qualquer vetor como a diferença de dois vetores não-negativos.

Assim sendo, considere-se a condição $Ax \leq b$. Esta condição é equivalente a $Ax^+ - Ax^- \leq b$, com $x^+, x^- \geq 0$. Para mais, $cx = cx^+ - cx^-$, pelo que isto sugere a consideração de transformar o pol

$$P = \begin{cases} \max_x cx \\ Ax \leq b \end{cases}$$

em

$$Q = \begin{cases} \max_{(x_+, x_-)} cx_+ - cx_- \\ Ax_+ - Ax_- \leq b \\ x_+, x_- \geq 0 \end{cases}$$

(Aqui, (x_+, x_-) representa apenas um vetor, de dimensão igual ao dobro da de x . O $+$ e $-$ em subscripto é suposto ser sugestivo de que um deles é a parte positiva e o outro a parte negativa de x .)

Repare-se que este segundo pol não tem a mesma dimensão que o primeiro: se no primeiro, o vetor x pertence a \mathbb{R}^n , no segundo, o vetor (x_+, x_-) pertence a \mathbb{R}^{2n} . Portanto, não é imediatamente óbvio que estes são ‘o mesmo’ pol, pelo que vamos fazer a seguinte definição:

Definição 2. Dados dois problemas de otimização linear, P e Q , dizemos que Q é uma f -tradução de P se se for possível resolver P sabendo a resposta de Q , usando f .

Em termos mais precisos, se f é uma função sobrejetiva $X_Q \rightarrow X_P$ tal que $q \in X_Q$ tem uma pontuação melhor que $q' \in X_Q$ (menor se Q é de minimização, maior se maximização) sse $f(q)$ tem melhor pontuação que $f(q')$.

Esta definição tem interesse devido à seguinte proposição:

Prop 1. Se Q é uma f -tradução de P , $f(S_Q) = S_P$.

Dem. Basta reparar que, por definição, $x \in S_Q$ sse x tem pontuação melhor ou igual que y para todo $y \in X_Q$, sse $f(x)$ tem pontuação melhor que $f(y)$ para todo $y \in X_Q$. Queremos então mostrar que isto é sse $f(x) \in S_P$.

Como f é sobrejetivo, isto implica que $f(x)$ tem pontuação melhor ou igual que todo z em X_P , o que por definição implica $f(x) \in S_P$. Pelo outro lado, se houver $y \in X_Q$ tal que $f(y)$ tem pontuação melhor que $f(x)$, então $f(x)$ claramente não pode ser solução, pois $f(y)$ é admissível de pontuação melhor que ele. \square

Isto mostra que se resolvermos uma tradução de um pol, é fácil recuperar a solução e conjunto admissível do problema original.

O caso anterior, em que $X_P = X_Q$ e $S_P = S_Q$ é o caso trivial em que f é a identidade. Neste caso, em que pretendemos adicionar a condição de positividade, é ligeiramente mais complicado.

Vamos mostrar que, para o P e Q definidos em cima, Q é uma f -tradução de P , em que f é a função $f(x_+, x_-) = x_+ - x_-$.

Dem. Primeiro, mostre-se que se um vetor pertence a X_Q , a sua imagem pertence a X_P .

Repare-se que, se $(x_+, x_-) \in X_Q$, então a sua imagem é $f(x_+, x_-) = x_+ - x_-$. Por definição, estes satisfazem $A(x_+ - x_-) \leq b$, donde $Af(x) \leq b$, que é a condição para $f(x) \in X_P$.

Para mostrar sobrejetividade, repare-se que se $x \in X_P$, tem-se $(x^+, x^-) \in X_Q$, e $f(x^+, x^-) = x^+ - x^- = x$.

Agora, queremos mostrar que (x_+, x_-) tem pontuação melhor que (y_+, y_-) sse $f(x_+, x_-)$ tem pontuação melhor que $f(y_+, y_-)$. Mas isto é trivial, pois a pontuação de $f(x_+, x_-)$ é a mesma que a de (x_+, x_-) , e o objetivo é, tanto em P como em Q , maximizar. \square

Assim sendo, podemos sempre assumir, sem perda de generalidade, que qualquer problema que tenhamos está na chamada *forma canónica*, isto é, é da forma

$$\begin{cases} \max_x cx \\ Ax \leq b \\ x \geq 0 \end{cases}$$

Repare-se que a nossa função de tradução f é linear. Isto vai ser relevante eventualmente. Dizemos que qualquer pol pode ser *traduzido linearmente para um pol na forma canónica*.

3.2.3 Desigualdades para igualdades (Forma padrão)

Outra forma útil de expressar um problema de otimização linear é a chamada forma padrão, da forma seguinte:

$$\begin{cases} \min_x cx \\ Ax = b \\ x \geq 0 \end{cases}$$

Já sabemos que podemos supor que os nossos problemas estão na forma canónica, sem perda de generalidade. Portanto, considere-se o pol

$$P = \begin{cases} \max_x cx \\ Ax \leq b \\ x \geq 0 \end{cases}$$

E vamos reescrevê-lo na forma padrão.

Para fazer a tradução, é preciso representar desigualdades com igualdades. Para isso, repare-se que $a \leq b$ é a mesma coisa que dizer que $a + y = b$, para algum y não-negativo. Assim, adicionamos as chamadas *variáveis de folga*: Considere-se o pol

$$Q = \begin{cases} \min_{(x,y)} -cx \\ Ax + y = b \\ x, y \geq 0 \end{cases}$$

Vamos mostrar que Q é uma f -tradução de P , com $f(x, y) = x$.

Dem. Primeiro, verifique-se que se $(x, y) \in X_Q$ se tem $f(x, y) \in X_P$. Repare-se que a condição $f(x, y) \geq 0$ é trivial, pois $x \geq 0$ por hipótese, e a condição $Af(x, y) \leq b$ é verdadeira pois $Af(x, y) = Ax \leq Ax + y = b$.

De seguida, é preciso verificar sobrejetividade. Suponhamos que se quer um objeto cuja imagem sob f seja $x \in X_P$. Considere-se, então, o vetor (x, y) , com $y = b - Ax$. Repare-se que este vetor pertence a X_Q , e a sua imagem é precisamente x , como queríamos demonstrar.

Finalmente, provemos que (x, y) tem melhor pontuação que (x', y') em Q sse x tem melhor pontuação que x' em P .

Mas isto é trivial, pois (x, y) tem melhor pontuação que (x', y') sse $-cx < -cx'$ sse $cx > cx'$, como queríamos demonstrar. \square

Isto mostra, então, que qualquer pol tem uma tradução linear para um pol na forma padrão.

3.2.4 Padrão para Canónico

Viu-se agora uma forma para traduzir problemas da forma canónica para a forma padrão. É possível, também, passar da forma padrão para a canónica (provou-se, aliás, que é possível passar de qualquer pol para a forma canónica) usando os procedimentos usados antes.

No entanto, é muito mais económico fazer o processo simples de transformar todas as igualdades em desigualdades do tipo \leq . Ou seja, passar de:

$$\begin{cases} \min_x cx \\ Ax = b \\ x \geq 0 \end{cases}$$

Para

$$\begin{cases} \max_x -cx \\ Ax \leq b \\ -Ax \leq -b \\ x \geq 0 \end{cases}$$

Já vimos antes que estas traduções são válidas.

4 Noções Topológicas

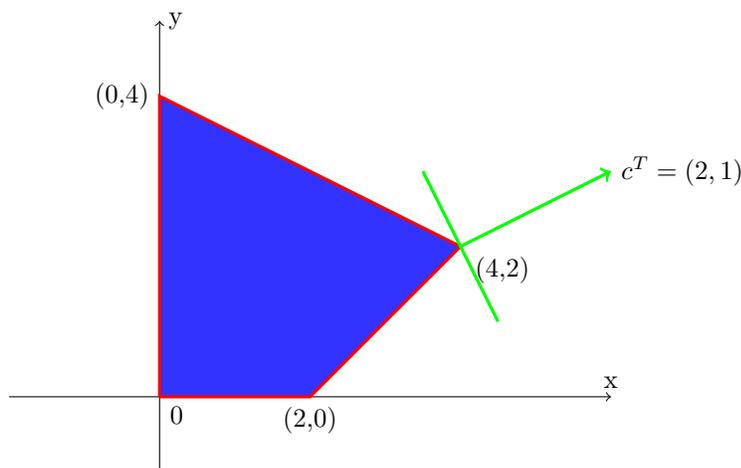
4.1 Intuição geométrica

Esta secção é dedicada a dar intuição geométrica para motivar os teoremas que se seguem. Vamos, a título de exemplo, partir de um pol específico a duas dimensões na forma canónica.

Considere-se o pol

$$P = \begin{cases} \max_{(x,y)} 2x + y \\ x + 2y \leq 8 \\ x - y \leq 2 \\ x, y \geq 0 \end{cases}$$

Vamos representar, no plano cartesiano, a região admissível X_P e o vetor $c^T = (2, 1)$.



Uma interpretação para o conjunto solução S_P é o conjunto de vetores admissíveis que maximiza o produto escalar com c^T . Intuitivamente, os vetores que ‘estão o mais possível na direção de c^T ’.

É possível ver (isto em termos intuitivos, claro) que não há nenhum vetor admissível ‘para lá’ da linha verde desenhada, o que indica que o conjunto solução será o conjunto $\{(4, 2)\}$.

Neste caso, o conjunto solução tem apenas um elemento, mas é concebível uma situação em que tivesse mais. Por exemplo, se o vetor c^T estivesse perpendicular ao lado $(0, 4) - (4, 2)$.

Isto é só especulação com base num desenho, mas dita especulação pode ser útil para motivar conclusões.

Por exemplo, com base nestas considerações, seria de esperar que, por exemplo, em geral o conjunto solução estará contido na fronteira do conjunto admissível, o que quer que isso signifique (a área marcada a vermelho na imagem). Também é fácil uma pessoa convencer-se que tem necessariamente de conter pelo menos um dos vértices – novamente, o que quer que isso signifique.

Seria também de esperar que no caso de o conjunto admissível ser limitado e não vazio, existe, necessariamente, solução. E, já agora, que o conjunto admissível é convexo.

Todos estes factos aleatórios serão provados, e mais, e alguns deles provar-se-ão úteis na procura de formas de resolver problemas de otimização.

4.2 Factos introdutórios

Recorde-se o leitor da definição de interior, fronteira e exterior de conjunto, e da definição de conjunto aberto e fechado.

Prop 2. *Fixo um pol P , X_P e S_P são fechados.*

Dem. Primeiramente, mostre-se que X_P é fechado.

Recorde-se que interseções finitas de fechados são fechadas. De seguida, note-se que, dado um pol P (que podemos supor sem perda de generalidade

estar na forma semicanónica) o seu conjunto admissível é o conjunto de $x \in \mathbb{R}^n$ que satisfazem as condições

$$a_i x \leq b_i \text{ para } i = 1, 2, \dots, m$$

Onde os a_i são vetores-linha, e os b_i são escalares.

Ou seja, X_P é a interseção

$$\bigcap_{i=1}^m \{x \in \mathbb{R}^n \mid a_i x \leq b_i\}$$

Repare-se agora que $\{x \in \mathbb{R}^n \mid a_i x \leq b_i\}$ é a pré-imagem sob uma função linear (e então contínua) de um conjunto fechado $(] - \infty, b_i])$, pelo que usando Cálculo II se conclui que estes conjuntos são todos fechados, pelo que X_P é fechado.

Para mostrar que S_P é fechado, repare-se que há dois casos: ou S_P é vazio, e então fechado; ou contém pelo menos um elemento s . Neste último caso, $S_P = X_P \cap \{x \mid c^T x \geq c^T s\}$ por definição, e isto é a interseção de dois conjuntos fechados, e então fechado. \square

Recorde-se do teorema de Weierstrass, que deverá ter dado em Cálculo II: uma função contínua definida num subconjunto fechado, limitado e não-vazio de \mathbb{R}^n (compacto não-vazio) tem máximo e mínimo. Assim sendo, conseguimos facilmente o seguinte corolário:

Prop 3. *Fixo um problema de otimização linear P , se X_P é limitado e não-vazio, S_P é não-vazio.*

Dem. Basta reparar que se X_P é limitado, por ser fechado (ver acima), é compacto. Assim sendo, como a função objetivo é linear e então contínua, tem valor máximo algures em X_P . Os valores maximizantes formam o conjunto solução. \square

Mais um detalhe topológico simples, antes de avançarmos para coisas mais específicas: denomine-se o interior de um conjunto X por $\text{int } X$, e a sua fronteira por ∂X .

Prop 4. *Se P é um pol de função objetivo não-nula, $S_P \subseteq \partial X_P$.*

Dem. Basta reparar que, se $x \in \text{int } X_P$, então há uma vizinhança ε de x contida em X_P . A ideia é que, então, se c^T não é o vetor nulo, podemos ‘avançar um bocadinho na direção de c^T ’. Em termos mais precisos, o vetor $y = x + \frac{\varepsilon}{2} \frac{c^T}{|c^T|}$ pertence a X_P , por distar de x menos de ε . Mas este vetor é ‘melhor’ que x , pois $c^T y = c^T x + c^T \frac{\varepsilon}{2} \frac{c^T}{|c^T|} = c^T x + \frac{\varepsilon |c^T|}{2} > c^T x$. Logo, nenhum $x \in \text{int } X_P$ pertence a S_P , donde S_P tem que estar todo contido na fronteira. \square

Só um último detalhe, antes de avançarmos para outras pastagens: uma caracterização do interior de X_P para um pol P , que se assume spdq estar na forma semicanónica.

Prop 5. *Seja P o pol*

$$P = \begin{cases} \max_x cx \\ Ax \leq b \end{cases}$$

e assumamos que nenhuma linha da matriz A é composta somente de zeros. (Isto é denotado por A é uma matriz própria)

Então, $\text{int } X_P$ é o conjunto de $x \in \mathbb{R}^n$ tal que $Ax < b$. Consequentemente, ∂X_P é o conjunto de x tal que $Ax \leq b$ e há pelo menos uma igualdade.

Dem. Relembre-se da seguinte identidade: $\text{int}(A \cap B) = \text{int } A \cap \text{int } B$. Mais geralmente, o interior de uma interseção finita é a interseção dos interiores.

Como tal, sendo que

$$X_P = \bigcap_{i=1}^m \{x \in \mathbb{R}^n \mid a_i x \leq b_i\}$$

onde a_i representa a i -ésima linha da matriz A ,

Para determinar $\text{int } X_P$ basta determinar os interiores destes conjuntos. Afirmamos que o interior do i -ésimo conjunto é

$$\text{int}\{x \in \mathbb{R}^n \mid a_i x \leq b_i\} = \{x \in \mathbb{R}^n \mid a_i x < b_i\}$$

Para justificar a inclusão \supseteq , note-se que o lado direito é um conjunto aberto contido em $\{x \in \mathbb{R}^n \mid a_i x \leq b_i\}$, e então contido no seu interior.

Para justificar a inclusão \subseteq , suponha-se que x pertence ao lado esquerdo. Mostrar-se-á que pertence ao lado direito.

Por pertencer ao interior daquele conjunto, sabemos que, para y numa vizinhança ε de x , $a_i y \leq b_i$. Considere-se, então, $y = x + \frac{\varepsilon}{2} \frac{a_i^T}{|a_i^T|}$. Por hipótese, $a_i y \leq b_i$. Mas tem-se também $a_i y = a_i x + a_i \frac{\varepsilon}{2} \frac{a_i^T}{|a_i^T|} = a_i x + \frac{\varepsilon |a_i^T|}{2} > a_i x$, donde se conclui $a_i x < a_i y \leq b_i$. Logo, x pertence ao lado direito. \square

5 Álgebra Linear

5.1 Intuição geométrica

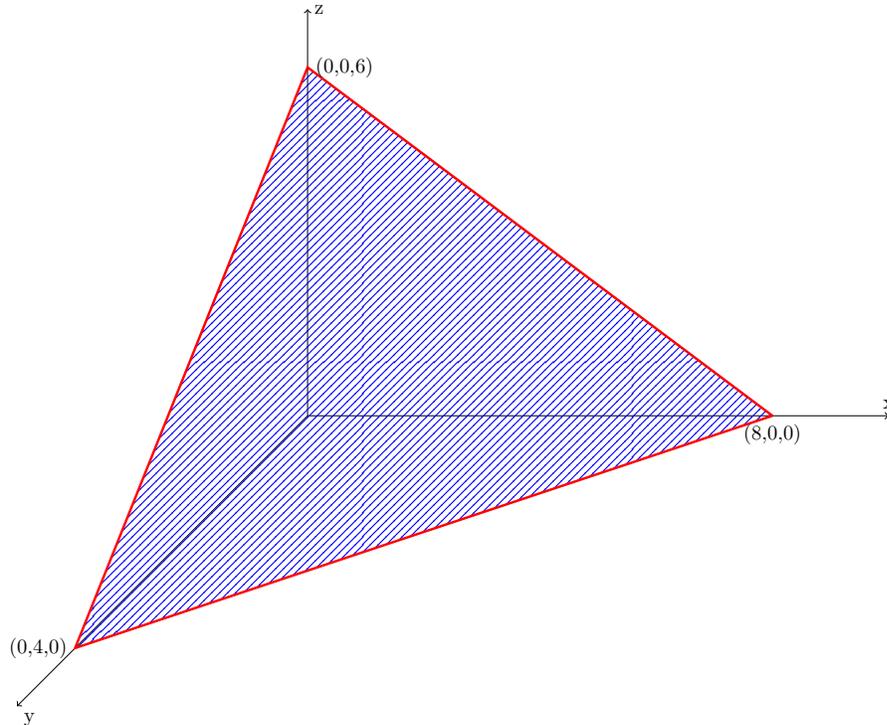
Queremos agora investigar a intuição que o conjunto solução contém sempre um vértice, no caso de este ser não-vazio.

Para este estudo, no entanto, é útil considerar, em vez do pol na forma canónica, o pol na forma padrão, visto que é mais fácil caracterizar o significado de ‘vértice’ neste contexto.

Considere-se o seguinte pol:

$$P = \begin{cases} \min_{(x,y,z)} 2x + y + 5z \\ 3x + 6y + 4z = 24 \\ x, y, z \geq 0 \end{cases}$$

O desenho do conjunto admissível é o seguinte:



Temos agora de definir o conceito de vértices no contexto de otimização linear, mas a figura deve tornar evidente que se tratam dos pontos com o maior número de coordenadas 0.

Vamos primeiro tentar definir o que queremos dizer por ‘vértice’.

5.2 Geometria

Dado o nosso contexto, é possível atribuir significado aos termos ‘vértice’, ‘aresta’, ‘lado’ e assim por diante. No entanto, por simplicidade, vamos apenas definir o significado de vértice.

Relembre-se da definição de plano em Álgebra Linear. Um plano em \mathbb{R}^n é um conjunto da forma $U = a + V$, em que $a \in \mathbb{R}^n$ e V é um subespaço vetorial de \mathbb{R}^n . Dizemos que U é um plano de dimensão m se V tiver dimensão m .

Um *hiperplano em \mathbb{R}^n* é um plano de dimensão $n - 1$.

Relembre-se que os planos podem também ser caracterizados como conjuntos-solução de equações lineares $Ax = b$ a n dimensões. A dimensão do plano é, então, a dimensão do núcleo da matriz A . Relembre-se também que, se um plano $a + V$ é definido por $Ax = b$, as colunas de A^T geram o espaço V^\perp .

Uma particularidade dos hiperplanos é que o seu espaço ortogonal tem dimensão 1. Assim sendo, um hiperplano pode sempre ser expresso da forma $w \cdot x = b$, para algum vetor não-nulo w e escalar b . Para mais, quaisquer duas

expressões que representem o mesmo hiperplano têm vetores colineares.

Isso permite fazer a seguinte definição:

Definição 3. Dado um plano H definido por $w \cdot x = b$, dizemos que este separa o espaço em dois semiespaços: H^+ , definido por $w \cdot x > b$ e H^- , definido por $w \cdot x < b$.

Devido à observação anterior, esta definição não depende do w escolhido, a menos de sinal. Ou seja, H^+ e H^- não dependem de w , exceto para distinguir qual é qual.

Diz-se que H separa o espaço em H^+ e H^- .

Estamos agora prontos para caracterizar o que queremos dizer com ‘vértice’.

Definição 4. Dado um conjunto $S \subseteq \mathbb{R}^n$, dizemos que $v \in S$ é um *vértice exterior de S* se existem $w \in \mathbb{R}^n$ e $b \in \mathbb{R}$ tal que $w \cdot v = b$ e $w \cdot x < b$ para todo $x \in S \setminus \{v\}$.

Esta definição é equivalente à seguinte:

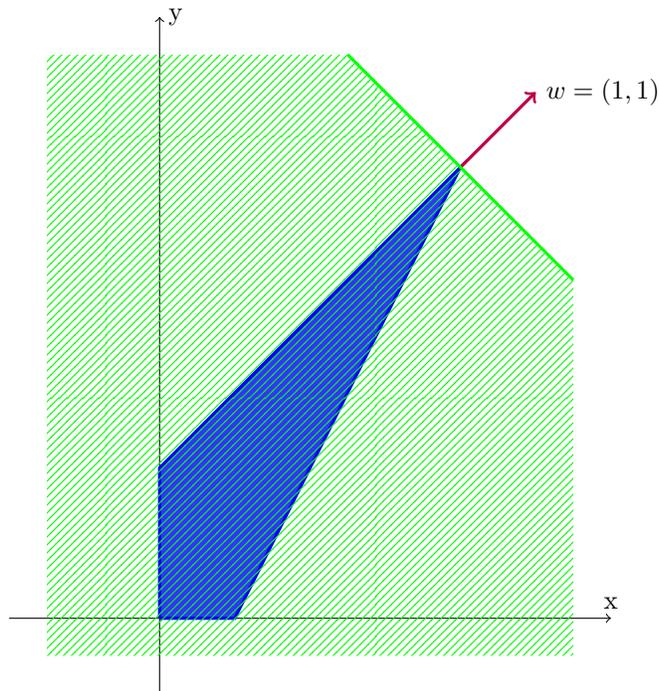
Definição. Nas mesmas condições, dizemos que $v \in S$ é vértice exterior se existe um hiperplano que separa v do resto do conjunto. Ou seja, se existe um hiperplano H que separe o espaço em H^+ e H^- tal que $v \in H$ e $S \setminus \{v\} \subseteq H^-$.

Repare-se que é óbvio que esta segunda definição implica a primeira, devido à forma como se definiram os subespaços. No entanto, a outra implicação não é tão óbvia, pois repare-se que não é exigido do vetor w que seja diferente de zero. Este detalhe tem pequena relevância, mas permitir-nos-á simplificar uma prova algures no futuro. Claro que se o vetor w for 0, o conjunto S contém apenas o ponto v , e portanto a segunda definição aplica-se com qualquer hiperplano que passe por v . Isto mostra que as duas definições são equivalentes.

Para exemplificar esta definição, considere-se o conjunto:

$$S = \{ (x, y) \mid x, y \geq 0, y \leq 2 + x, y \geq -2 + 2x \}$$

Em baixo, está uma representação da afirmação: “o ponto $(4, 6)$ é vértice de S , pois pondo $w = (1, 1)$ e $b = 10$, tem-se $w \cdot (4, 6) = b$ e para todos os pontos (x, y) em $S \setminus \{(4, 6)\}$ tem-se $w \cdot (x, y) < b$ ”.



A reta verde representa o hiperplano (reta) H definido por $w \cdot (x, y) = 10$, e a área pintada a verde representa o semiespaço H^- definido por $w \cdot (x, y) < 10$. Repare-se que todo o conjunto S , exceto o ponto $(4, 6)$, está contido neste semiespaço.

5.3 Vértices de conjuntos admissíveis

Estamos agora prontos para tentar caracterizar os vértices de conjuntos admissíveis do pol na forma padrão. A seguinte proposição ser-nos-á útil:

Prop 6. *Seja S um conjunto em \mathbb{R}^n . As seguintes quatro afirmações não podem ser todas verdade ao mesmo tempo:*

- v é vértice de S
- $v + h \in S$
- $v - h \in S$
- $h \neq 0$

Dem. Para mostrar isto, vamos supor que três destes são verdade, e mostramos que o quarto tem que ser falso. Nomeadamente, suponhamos que v é vértice de S , $v + h \in S$ e $h \neq 0$.

Assim sendo, visto que v é vértice, existem w e b tal que $w \cdot v = b$ e $w \cdot x < b$ para todo $x \in S$ diferente de v . Em particular, $w \cdot (v + h) < b$. Mas então

$w \cdot h < 0$, donde $w \cdot (v - h) > b$, e então concluímos que $v - h$ não pode pertencer a S . \square

Na caracterização de vértices é-nos útil falar das coordenadas positivas de vetores, e das colunas correspondentes da matriz A . (A razão para isto será evidente em breve.) Assim sendo, adotamos as seguintes notações:

Notação. No nosso contexto, considere-se o seguinte pol a n dimensões:

$$\begin{cases} \min_x cx \\ Ax = b \\ x \geq 0 \end{cases}$$

Em que x é um vetor em \mathbb{R}^n , A é uma matriz $m \times n$ e b é um vetor em \mathbb{R}^m . Por conveniência, seja N o conjunto $\{1, 2, \dots, n\}$.

Fixo um vetor x :

Usamos P_x para denotar o conjunto de índices i tal que $x_i > 0$.

Dado um conjunto de índices $B \subseteq N$, x_B representa o vetor em $\mathbb{R}^{\#B}$ obtido de x resultante apenas das coordenadas em B . Por exemplo, se B é o conjunto $\{2, 5, 9\}$, x_B seria o vetor em \mathbb{R}^3 dado por (x_2, x_5, x_9) . Análogamente, A_B denota a submatriz de A considerando apenas as colunas com índices em B .

Antes de fazer a prova ‘a sério’, será feito um esboço de prova para suportar a afirmação feita há bocado: “os vértices tratam-se dos pontos com o maior número de coordenadas 0”. Este esboço de prova vai-nos servir para motivar a caracterização (completa) dos vértices de X_P , para P na forma padrão.

Seja k a característica da matriz A . Recorde-se que a característica de uma matriz equivale à dimensão do seu espaço de colunas. Afirmamos que qualquer vértice de X_P tem no máximo k coordenadas maiores que zero.

Para justificar isto, considere-se um $x \in X_P$ tal que $\#P_x > k$. Queremos justificar que este não pode ser um vértice, e vamos fazer isso usando a proposição 6.

Nomeadamente, vamos arranjar um h diferente de zero tal que $x + h$ e $x - h$ pertencem a X_P , o que, devido a esta proposição, implica que x não é vértice de X_P . Para arranjar este h , repare-se nas condições que ele tem que obedecer:

- $h \neq 0$
- $Ah = 0$
- $x \pm h \geq 0$

Em particular, a última condição diz-nos algo importante: para todo i , se $x_i = 0$, h_i tem que ser 0. Caso contrário, um daqueles dois vetores teria a i -ésima casa negativa.

Assim sendo, isso limita as nossas escolhas de h . Este tem necessariamente as casas em $N \setminus P_x$ igual a zero, pelo que resta definir h_{P_x} .

Repare-se, agora, que, sob estas condições, $Ah = A_{P_x} h_{P_x}$, e visto que $\#P_x$ é maior do que a característica de A , a matriz A_{P_x} é singular, e então existe um h diferente de zero (chamemos-lhe \tilde{h}) que satisfaz $A\tilde{h} = 0$ e \tilde{h}_i é igual a zero para $i \notin P_x$. Falta apenas assegurar que $x \pm \tilde{h} \geq 0$, mas isto pode não ser verdade.

Para o passo final, repare-se que estas duas condições ($A\tilde{h} = 0, \tilde{h} \neq 0$) não deixam de ser verdade se multiplicarmos \tilde{h} por um escalar t diferente de zero.

A ideia é a seguinte: para todo $i \in N$ tem-se um dos dois casos: ou $x_i = 0$ e então, como $\tilde{h}_i = 0$, $x_i + t\tilde{h}_i \geq 0$ independentemente de t , ou $x_i > 0$ e então $x_i + t\tilde{h}_i \geq 0$ para qualquer t de módulo pequeno o suficiente. Considere-se o m mínimo destes módulos, e tem-se que, para t de módulo menor que m , $x \pm t\tilde{h}_i \geq 0$. Pondo $h = t\tilde{h}_i$, obtemos o resultado desejado: $x \pm h \in X_P$ e $h \neq 0$. Assim sendo, x não pode ser vértice de X_P .

Repare-se que neste esboço de prova, a condução que $\#P_x > k$ foi apenas usada para justificar que as colunas de A_{P_x} são linearmente dependentes. Acontece que esta nova condição, mais fraca, não é só suficiente como necessária. Provamos, agora, a

Prop 7. *Seja P o seguinte pol na forma padrão:*

$$\begin{cases} \min_x cx \\ Ax = b \\ x \geq 0 \end{cases}$$

Então, $x \in X_P$ é vértice de X_P sse A_{P_x} tem as suas colunas todas linearmente independentes.

Dem. Primeiro, a parte que já está feita: se as colunas de A_{P_x} não forem linearmente independentes, x não é vértice. (\rightarrow) Construa-se $\tilde{h} \neq 0$ tal que $\tilde{h}_i = 0$ para $i \notin P_x$, e \tilde{h}_{P_x} é tal que $A_{P_x} \tilde{h}_{P_x} = 0$. Isto pode ser feito porque, por hipótese, a equação $A_{P_x} h = 0$ tem soluções não-triviais.

Considere-se a função $f(t) = x + t\tilde{h}$. Pretendemos arranjar t tal que $f(t), f(-t) \geq 0$, pois assim ter-se-ia que $x \pm t\tilde{h} \in X_P$ com $t\tilde{h} \neq 0$, o que justifica x não ser vértice conforme a prop 6.

Para este objetivo, decomponha-se esta função em $f_{P_x}(t)$ e $f_{N \setminus P_x}(t)$. A segunda é sempre ≥ 0 (por ser sempre igual a zero), pelo que basta examinar a primeira.

Mas repare-se que, como f é contínua (soma de uma constante com uma função linear), a sua restrição às coordenadas de P_x também. Para mais, note-se que $f_{P_x}(0) > 0$, pelo que, por continuidade, existe uma vizinhança ε de zero onde $f_{P_x}(t) > 0$, e então $f(t) \geq 0$. Escolhendo um t diferente de zero nesta vizinhança, por exemplo, $t = \varepsilon/2$, temos $f(t), f(-t) > 0$, como desejado.

A partir daqui, aplicando o raciocínio anterior, usando $h = t\tilde{h}$, concluímos que x não é vértice de X_P .

Agora, a segunda parte (\leftarrow): mostrar que se $x \in X_P$ e as colunas de A_{P_x} são linearmente independentes, então x é vértice.

Para fazer isto, considere-se o vetor w definido como

$$w_i = (0 \text{ se } i \in P_x, -1 \text{ caso contrário})$$

E o escalar 0.

Então, claramente temos $w \cdot x = 0$.

Agora, mostre-se que, para todo $y \in X_P \setminus \{x\}$ se tem $w \cdot y < 0$. Isto é claramente equivalente (dado que $y \in X_P$ e então $y \geq 0$) a mostrar que $y_i \neq 0$ para algum $i \notin P_x$. Ou seja, que $P_y \not\subseteq P_x$.

Para este efeito, suponha-se que $y \in X_P$ e $P_y \subseteq P_x$. Vamos mostrar que $y = x$.

Se $y \in X_P$, temos $Ay = b$. Mas como $P_y \subseteq P_x$, temos que $Ay = A_{P_x} y_{P_x} = b$. Mas como as colunas de A_{P_x} são linearmente independentes, a solução do sistema $A_{P_x} v = b$, se existir (neste caso existe), é única. Assim sendo, $y_{P_x} = x_{P_x}$, e como as outras coordenadas de y são todas 0, como as de x , temos $y = x$, como queríamos demonstrar.

Logo, para $y \in X_P \setminus \{x\}$ tem-se sempre $w \cdot y < 0$, como se pretendia demonstrar, e então x é vértice de X_P . \square

Estamos agora em condições de justificar o seguinte:

Prop 8. *Seja P o pol*

$$\begin{cases} \min_x cx \\ Ax = b \\ x \geq 0 \end{cases}$$

Se a função $x \mapsto cx$ é minorada em X_P , para todo $v \in X_P$ existe um vértice z de X_P tal que $cz \leq cv$.

Dem. Se $v \in X_P$, o conjunto $\{x \in X_P \mid cx \leq cv\}$ é não-vazio. A função $x \mapsto \#P_x$ vai deste conjunto para \mathbb{N}_0 , pelo que é minimizada nalgum ponto. Chamemos-lhe z . Mostraremos que z é vértice de X_P .

Para fazer isto, usaremos a caracterização feita há pouco: mostraremos que as colunas de A_{P_z} são linearmente independentes, o que justifica que z é vértice.

A prova faz-se por contrarecíproco. Para mostrar que um z que minimize $\#P_z$ é vértice, supõe-se que não é vértice e mostra-se que não minimiza $\#P_z$.

Suponha-se, então, que z não é vértice, ou seja, que as colunas de A_{P_z} são linearmente dependentes. Vamos arranjar um elemento \tilde{z} de S_P tal que $\#P_{\tilde{z}} < \#P_z$.

Sabendo que as colunas de A_{P_z} são linearmente dependentes, existe solução não-trivial do sistema $A_{P_z} x = 0$. Seja \tilde{h} tal que $\tilde{h}_i = 0$ para $i \notin P_z$, e \tilde{h}_{P_z} é solução não-trivial de $A_{P_z} \tilde{h}_{P_z} = 0$. Suponha-se, sem perda de generalidade, que $c\tilde{h} \geq 0$. Podemos fazer isto, pois, se não for o caso, substitua-se \tilde{h} por $-\tilde{h}$.

Repare-se que, como consequência, podemos também assumir que \tilde{h} tem pelo menos uma coordenada positiva. Isto pois, se $c\tilde{h} = 0$ podemos simplesmente trocar \tilde{h} por $-\tilde{h}$, e se $c\tilde{h} > 0$, este vetor tem de certeza pelo menos uma coordenada positiva. Caso contrário, todos os vetores da forma $x - t\tilde{h}$ seriam

admissíveis (verifique) e, fazendo $t \rightarrow +\infty$, temos que a função objetivo toma valores arbitrariamente pequenos em X_P , o que contraria o enunciado.

Assim sendo, temos as seguinte duas hipóteses sobre \tilde{h} : $c\tilde{h} \geq 0$ e \tilde{h} tem pelo menos uma coordenada positiva.

Assim sendo, para todo o t tal que $z - t\tilde{h}$ é admissível, este tem pontuação melhor ou igual do que v .

A ideia é, agora, escolher t de modo a que $z - t\tilde{h}$ tenha menos coordenadas positivas do que z , mas continue admissível. Para fazer isto, efetivamente, o que se faz é ‘subir o valor de t continuamente até uma das coordenadas ser zero’. (Isto vai ser formalizado já a seguir.) Essa coordenada ficará zero, e nenhuma coordenada que fosse antes zero deixou de o ser, pelo que o número de coordenadas positivas diminuiu... E o vetor continua a ser admissível porque a condição $Ax = b$ nunca deixa de ser verdade, e a condição $x \geq 0$ continua a ser verdade porque não deixamos nenhuma coordenada ir para os negativos.

Em termos formais: considere-se o conjunto $T = \{\frac{z_i}{\tilde{h}_i} \mid i \in P_{\tilde{h}}\}$. Este conjunto é não-vazio (assumimos que \tilde{h} tinha pelo menos uma coordenada positiva), finito e todos os seus elementos são positivos (pois $P_{\tilde{h}} \subseteq P_z$). Considere-se, então, o seu mínimo, α . Vamos ver que o vetor $\tilde{z} = z - \alpha\tilde{h}$ é admissível e tem menos coordenadas positivas do que z .

Como já vimos, obedece à condição $A\tilde{z} = b$, pelo que basta certificarmos-nos que $\tilde{z} \geq 0$. Fazemos isto coordenada a coordenada.

Examine-se $\tilde{z}_i = z_i - \alpha\tilde{h}_i$. Se $i \notin P_{\tilde{h}}$, claramente $\tilde{z}_i \geq z_i \geq 0$. Pelo outro lado, se $i \in P_{\tilde{h}}$, temos que $\alpha \leq z_i/\tilde{h}_i$, e então $-\alpha \geq -z_i/\tilde{h}_i$, donde $z_i - \alpha\tilde{h}_i \geq z_i - \frac{z_i}{\tilde{h}_i}\tilde{h}_i = 0$. Como se pretendia demonstrar.

Assim sendo, o vetor \tilde{z} é admissível. Já vimos que tem pontuação melhor ou igual do que v . E, finalmente, mostramos que tem menos coordenadas positivas do que z .

Repare-se que para todo i , se $z_i = 0$ então $\tilde{h}_i = 0$, pelo que nenhuma coordenada positiva ‘novas’ aparecem. Pelo outro lado, pelo menos uma coordenada positiva fica nula. Nomeadamente, visto que α é o valor mínimo do conjunto $\{\frac{z_i}{\tilde{h}_i} \mid i \in P_{\tilde{h}}\}$, é da forma $\alpha = z_j/\tilde{h}_j$ para algum j . Vamos examinar a coordenada j de \tilde{z} .

Sabemos que $z_j > 0$. Mas também temos que $\tilde{z}_j = 0$, pois $\tilde{z}_j = z_j - \alpha\tilde{h}_j = z_j - \frac{z_j}{\tilde{h}_j}\tilde{h}_j = 0$. Isto mostra, então, que \tilde{z} tem pelo menos uma coordenada positiva a menos que z , o que mostra que \tilde{z} é o elemento que pretendíamos arranjar. Isto conclui a nossa prova. \square

Um corolário útil, cuja demonstração é deixada como exercício ao leitor:

Prop 9. *Se S_P é não-vazio, contém pelo menos um vértice de X_P .*

Outro corolário:

Prop 10. *Dado um pol P na forma padrão, se o conjunto admissível não for vazio, tem pelo menos um vértice.*

Dem. Já provámos que, dado um elemento v do conjunto admissível, existe um vértice de pontuação menor ou igual que v . Em particular, isto significa que existe um vértice. \square

5.4 O método dos vetores básicos

Estamos perto de ter um método primitivo de resolução de problemas de otimização linear. Dado um pol, podemos sempre pô-lo na forma padrão. Uma vez nesta forma, podemos examinar os vértices.

A ideia é que, em princípio (vamos provar isto) haverá um número finito de vértices. Se o problema tiver solução, o vértice de melhor pontuação será solução. Daí em diante, é preciso verificar que, de facto, S_P é não-vazio. Os detalhes de tal verificação serão feitos mais tarde: por agora, arranjaríamos um método de encontrar os vértices de X_P para P na forma padrão.

A ideia é a seguinte: considere-se um conjunto de índices $B \subseteq N$ tal que A_B tem colunas linearmente independentes. Então, a equação $A_B x = b$ tem zero ou uma soluções. Isto dá-nos o primeiro método, mais primitivo possível, de encontrar os vértices de X_P :

Para todo o $B \subseteq N$:

Se as colunas de A_B são linearmente independentes:

Resolver o sistema $A_B x = b$.

Se este tiver solução x :

Se $x \geq 0$:

Definir $v \in \mathbb{R}^n$ de modo a que $v_B = x$ e $v_{N \setminus B} = 0$

Adicionar v à lista de vértices.

Este método funciona, mas tem um enorme problema: se estamos em dimensão n , temos 2^n sistemas para resolver. Procuramos, então, uma forma mais computacionalmente eficiente de fazer as coisas.

Lembre-se da Álgebra Linear que qualquer conjunto linearmente independente de vetores pode ser extendido a uma base. Isso é refletido na seguinte proposição:

Prop 11. *Dada uma matriz A $m \times n$ e um conjunto de índices $B \subseteq N$, dizemos que B é uma base de índices se as colunas de A_B formam uma base de $\text{Im } A$.*

Se $B \subseteq N$ é tal que as colunas de A_B são linearmente independentes, existe uma base de índices $B' \supseteq B$.

Dem. Esta prova será feita mostrando que, se B é tal que A_B tem colunas linearmente independentes mas B não é base de índices, é possível adicionar um elemento b a B tal que as colunas de $B \cup \{b\}$ continuem a ser linearmente independentes. Feito isto, o algoritmo para extender B a uma base passa a ser repetir este processo até se ter um conjunto de tamanho igual à característica de A .

Assim sendo, basta mostrar este passo. Suponha-se que B está sob estas condições, mas não é uma base de índices. Então, tem de existir um vetor-coluna de A que não está em $\text{Im } A_B$. Isto pois, caso contrário, $\text{Im } A = \text{Im } A_B$, e B seria uma base de índices. Assim sendo, escolha-se uma tal coluna, e seja b o seu índice. As colunas de $A_{B \cup \{b\}}$ são linearmente independentes, como o leitor poderá facilmente verificar, o que termina a nossa prova. \square

Dizemos que um vetor x é básico (no contexto de um pol P na forma padrão) se P_x está contido numa base de índices B .

O que a proposição anterior mostra é que todo vértice é básico (pegue-se em P_x e extenda-se a uma base). Pelo outro lado, é fácil ver que qualquer vetor básico x tem a propriedade que as colunas de A_{P_x} são linearmente independentes. Temos, então, que os vetores admissíveis básicos e os vértices de X_P são os mesmos.

As bases de índices têm uma propriedade notável: no contexto do pol

$$P = \begin{cases} \min_x cx \\ Ax = b \\ x \geq 0 \end{cases}$$

Se a equação $Ax = b$ é impossível, este problema trivialmente não tem solução, visto que o conjunto admissível é vazio. Se isto acontece, dizemos que o pol P é *vazio*.

Pelo outro lado, se $b \in \text{Im } A$, dada uma base de índices B , temos que o sistema de equações $A_B x = b$ tem uma e só uma solução.

Outra observação importante é que as bases de índices de A têm todas o mesmo tamanho igual à dimensão do espaço de colunas de A , ou seja, a sua característica.

Assim sendo, considere-se o seguinte (novo) algoritmo.

Verificar se o problema é vazio.

Caso afirmativo, a lista de vértices é vazia.

Caso contrário:

Seja k a característica de A .

Para todo o $B \subseteq N$ tal que $\#B = k$:

Se as colunas de A_B são linearmente independentes:

Resolver o sistema $A_B x = b$.

Se $x \geq 0$:

Definir $v \in \mathbb{R}^n$ de modo a que $v_B = x$ e $v_{N \setminus B} = 0$

Adicionar v à lista de vértices.

Este algoritmo é muito mais económico do que o anterior. Ao passo que no outro havia 2^n possibilidades para verificar, aqui é preciso apenas verificar $\binom{n}{k}$ conjuntos, onde k é a característica de A .

Para obter um candidato a solução basta, agora, calcular os vértices de X_P , v_1, v_2, \dots, v_ℓ , calcular cv_i para $i = 1, \dots, \ell$ e ver qual destes é menor.

Pegando, então, num vértice v_j que minimize cv_j , se o conjunto solução for não-vazio, contém v_j . Precisamos, então, de um método de verificar se um vértice é solução. É a isso que nos dedicamos de seguida. Mas antes, algumas proposições extremamente úteis:

Prop 12. *Dado um pol P na forma padrão, se a função-objetivo for minorada e o conjunto admissível não for vazio, o pol tem solução.*

Dem. Já vimos que X_P tem pelo menos um vértice, e já vimos que tem um número finito. Considere-se então um vértice z de menor pontuação. Mostraremos que este é solução.

Fixe-se um ponto $x \in X_P$ arbitrário. Como a função-objetivo é minorada, sabemos que existe um vértice \tilde{z} de pontuação menor ou igual do que x , e sabemos que z tem pontuação menor ou igual do que \tilde{z} . Logo, a pontuação de z é menor igual do que a de x . Como x é arbitrário, z é solução. \square

Prop 13. *Dado um pol P de minimização, se a função-objetivo for minorada e o conjunto admissível não for vazio, o pol tem solução.*

Da mesma forma, se P for de maximização, a função-objetivo for majorada e o conjunto admissível não for vazio, o pol tem solução.

Dem. Dado um pol arbitrário, sabemos, pelos conteúdos da primeira secção, que podemos traduzi-lo linearmente para a forma padrão com uma certa função linear sobrejetiva f que é monótona para a pontuação.

Isto significa que se a função objetivo é minorada/majorada no pol original, também será minorada no pol traduzido. Isto pois repare que, para as traduções que fizemos, se começarmos com um pol de minimização, a pontuação de $f(x)$ é igual à pontuação de x . Caso contrário, é a sua negação. Assim sendo, se M é minorante da pontuação no pol traduzido, M será minorante, ou $-M$ será majorante, no original, consoante se este é de minimização ou maximização.

Logo, se a função-objetivo for majorada/minorada, será minorada no pol traduzido, pelo que este terá solução, pelo que o pol original tem solução. \square

6 Cones

6.1 Introdução

No estudo do pol na forma padrão e canónica, é muito comum ver expressões da forma Ax com $x \geq 0$. Ou seja, combinações lineares não-negativas das colunas de A .

Este fenómeno é tão comum que tem um nome.

Definição 5. Sejam a_1, a_2, \dots, a_r vetores de \mathbb{R}^n . O cone formado por a_1, \dots, a_r é o conjunto:

$$\left\{ \sum_{i=1}^r t_i a_i \mid t \in \mathbb{R}^r, t \geq 0 \right\}$$

Em particular, é muito comum estes vetores serem as colunas de uma matriz. Assim sendo, dada uma matriz A , definimos *o cone das colunas de A* , denominado $CC A$, como:

$$\{ Ax \mid x \in \mathbb{R}^n, x \geq 0 \}$$

É fácil verificar que isto é o mesmo que o cone formado pelas colunas de A .¹

Nesta secção, estudaremos os cones, que têm surpreendente aplicação ao problema de otimização linear. O nosso principal resultado será o chamado *lema de Farkas*. Este e as suas variantes serão de grande utilidade para encontrar métodos para descobrir se um vetor é solução.

Um destes métodos é o chamado *método das linhas ativas*, e vamos esboçar metade da sua prova de correção, com o intuito de motivar o que se segue.

6.2 Método das linhas ativas (Parte 1)

Suponhamos que temos um pol na forma semicanónica:

$$P = \begin{cases} \max_x cx \\ Ax \leq b \end{cases}$$

Já vimos que, para o pol na forma padrão, o conjunto-solução, sempre que não-vazio, contém pelo menos um vértice. Em geral, infelizmente, isto não é verdade. Por exemplo, o pol mais trivial possível a uma dimensão:

$$\begin{cases} \max_x 0x \\ (\text{Sem condições...}) \end{cases}$$

Este pol tem solução (qualquer x , na realidade) mas não tem nenhum vértice. Isto mostra que a proposição “o conjunto solução, se não-vazio, contém pelo menos um vértice” não é verdade em geral. No entanto, é possível arranjar resultados parecidos, em espírito.

Para arranjar uma medida de ‘quão na fronteira estamos’, fazemos a seguinte definição:

Definição 6. Dado o pol P definido acima, a i -ésima linha de A (denotada, por agora, de a_i) diz-se *ativa relativamente a x* se $a_i x = b$.

Já vimos no início que x pertence à fronteira de X_P sse A tem pelo menos uma linha ativa relativamente a x . O que vamos ver nesta secção é que é possível averiguar se um vetor é solução examinando as linhas ativas.

¹O leitor poderá perguntar-se o que acontece ao cone gerado por zero elementos. Neste caso, o nosso conjunto será o conjunto que contém apenas um elemento (\mathbb{R}^0 tem um elemento, e todos os seus índices são ≥ 0) e este será a soma vazia, que dá zero. Logo, o cone gerado pelo vazio é o conjunto $\{0\}$. Em termos matriciais, isto equivale ao cone gerado por uma matriz $n \times 0$. Tendo isto em conta, é fácil ver que o conjunto vazio *não* é um cone sob esta definição.

Definição 7. Dado um vetor x de X_P , A^x representa a matriz das linhas ativas de x em A .

Prop 14. *Seja x um vetor admissível de P . x não é solução sse existe $w \in \mathbb{R}^n$ tal que $cw > 0$ e $A^x w \leq 0$*

Dem. Prove-se primeiro a implicação (\leftarrow). Suponha-se que existe $w \in \mathbb{R}^n$ tal que $cw > 0$ e $A^x w \leq 0$. A ideia é ‘somar w em pequenas quantidades’, de modo a obter um vetor admissível de pontuação melhor do que x .

É óbvio que se existe t positivo tal que $x + tw \in X_P$, x não é solução, pois este novo vetor tem pontuação melhor do que x .

Para arranjar tal t , considere-se cada linha de $A(x + tw)$.

Para todo i , há dois casos:

Ou a_i é linha ativa, e então $a_i(x + tw) = a_i x + ta_i w \leq a_i x = b_i$ para qualquer t ,

Ou a_i não é linha ativa, e então $a_i(x + tw) = a_i x + ta_i w$. Como $a_i x < b_i$, temos que, para t de módulo menor que $|\frac{b_i - a_i x}{a_i w}|$, $a_i x + ta_i w \leq b_i$.

Considere-se então t pequeno o suficiente para que $a_i(x + tw) \leq b_i$ para todo i (é fácil ver que tal t existe) e temos, como desejado, que este vetor é admissível e tem pontuação melhor do que x .

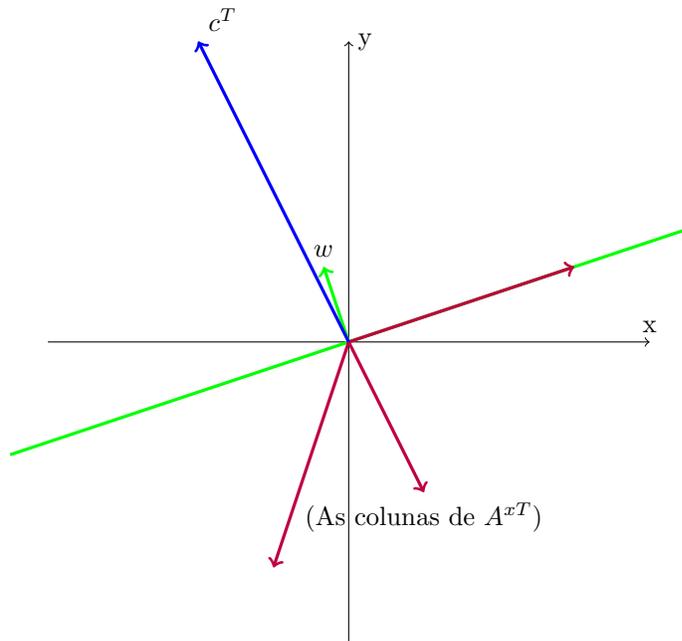
Para provar a implicação (\rightarrow) suponha-se que x não é solução. Então, existe um vetor $x + w$ de pontuação melhor do que x . Vamos mostrar que este w obedece às condições do enunciado.

Obviamente $cw > 0$, pois $c(x + w) = cx + cw > cx$ por hipótese. Falta então provar que $A^x w \leq 0$.

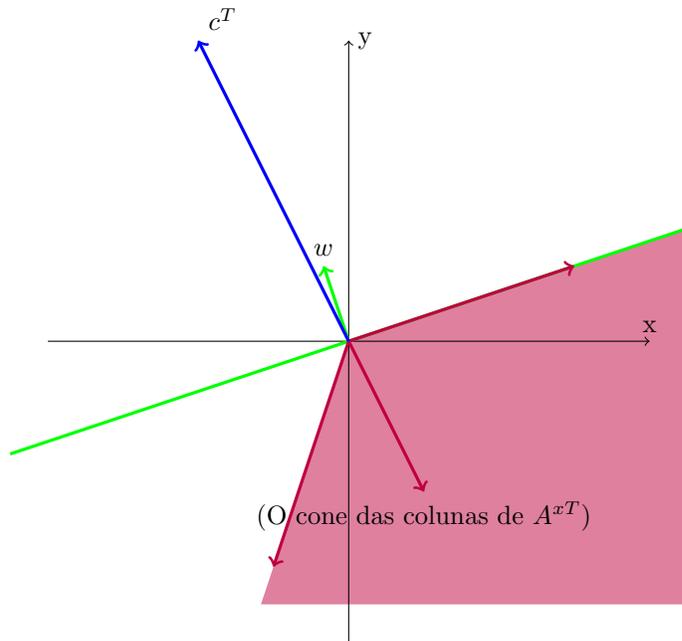
Mas se isso não fosse o caso, e houvesse i tal que a_i é linha ativa e $a_i w > 0$, ter-se-ia que a linha i de $A(x + w)$ seria $a_i x + a_i w > a_i x = b_i$, e então $x + w$ não seria admissível. Contradição. \square

Esta proposição vai-nos ajudar a desenvolver o método das linhas ativas. Falta, agora, caracterizar quando é que existe w tal que $cw > 0$ e $A^x w \leq 0$. Para dar uma ideia, visualize-se a situação.

Repare-se, primeiro que tudo, que a condição $cw > 0$ implica que w é diferente de zero, pelo que w define um hiperplano. Para mais, recorde-se que os hiperplanos dividem o espaço em dois. A condição que temos, então, é que, se considerarmos c^T e as colunas de A^{xT} como vetores em \mathbb{R}^n , o hiperplano definido por w separa c^T das colunas da matriz:



Vamos ver esta mesma imagem, mas, a vermelho, está o cone das colunas de A^{xT} .



A ideia é então a seguinte: mostrar que se c^T está fora deste cone, é possível arranjar tal w . Isto porque verificar se c^T está fora do cone é possivelmente mais fácil. Por exemplo, se as linhas de A^x são linearmente independentes, a equação

$A^T x = c$ tem no máximo uma solução. É então possível tentar encontrar tal solução. Se ela existir e for ≥ 0 , então c está no cone desejado, e x é solução. Caso contrário, não está, e x não é solução.

Isto motiva, então, o enunciado do lema de Farkas:

Conjetura 1. (*Lema de Farkas*)

Seja A uma matriz $n \times m$, v um vetor de \mathbb{R}^n .

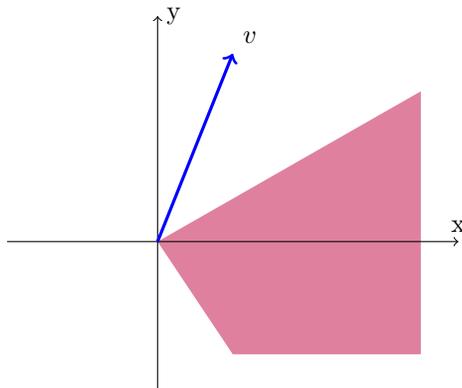
$v \notin \text{CCA}$ sse existe w tal que $w \cdot v > 0$ e $A^T w \leq 0$.

Focamo-nos, agora, em provar este lema.

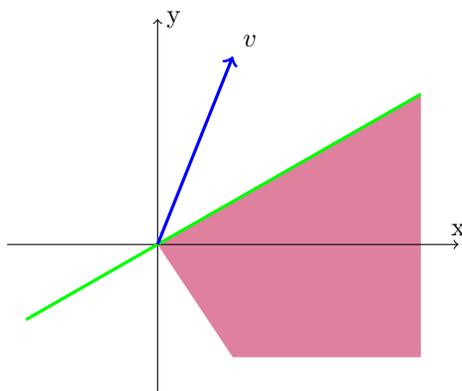
6.3 O Lema de Farkas (Parte 1)

Para provar o Lema de Farkas vai-nos ser necessário provar vários resultados intermédios. É, no entanto, feito o esboço da prova, para motivar esses resultados.

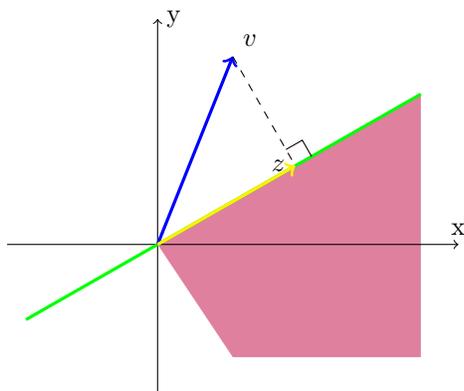
O objetivo é, dado um cone C e um ponto v fora deste:



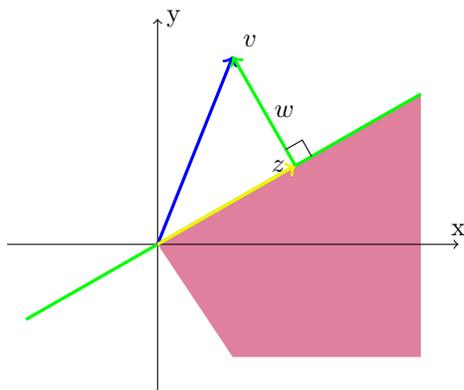
Encontrar um hiperplano que os separe.



O que vamos fazer é considerar a projeção ortogonal de v , digamos z , no cone



E consideramos o plano definido por $w = v - z$.



A partir daqui, há dois grandes passos que é preciso justificar:

- O que queremos dizer por projeção ortogonal, e como a encontramos?
- Como sabemos que o plano assim definido separa o ponto v do cone C ?

Dedicamo-nos agora ao primeiro.

6.4 Cones primitivos

Em geral, a ‘projeção ortogonal de x em Y ’ pode ser definida como o $y \in Y$ que minimize $|x - y|$. No entanto, é preciso assegurar que esta função pode ser minimizada.

Felizmente, em certas condições, a existência deste mínimo é-nos garantida.

Prop 15. *Se $x \in \mathbb{R}^n$ e Y é um subconjunto fechado e não-vazio de \mathbb{R}^n , existe $y \in Y$ tal que $|x - y| \leq |x - y'|$ para todo $y' \in Y$.*

Dem. Primeiro que tudo, considere-se o caso em que Y é limitado. Assim sendo, visto que também é fechado, é compacto. Logo, a função $f(y) = |x - y|$ tem mínimo por Weierstrass, o que mostra a nossa afirmação.

No caso de Y não ser limitado, pegue-se num y' qualquer em Y . Considere-se $Y' = Y \cap \{y \in \mathbb{R}^n \mid |x - y| \leq |x - y'|\}$. Este conjunto já é compacto, e então o raciocínio anterior aplica-se, e existe y que minimize f em Y' .

Mas se y minimiza f em Y' , também minimiza f em Y , pois para qualquer y'' em $Y \setminus Y'$ tem-se $f(y) \leq f(y') \leq f(y'')$. \square

Assim sendo, dado um ponto $x \in \mathbb{R}^n$ e um conjunto fechado Y , sabemos que existe a chamada *projeção ortogonal de x em Y* , ou seja, o $y \in Y$ que minimiza $|x - y|$.

Precisamos, agora, de assegurar que os cones são fechados.

Repare-se que qualquer cone em \mathbb{R}^n pode ser expresso como $\{Ax \mid x \in \mathbb{R}^m, x \geq 0\}$ para alguma matriz $m \times n$. (As suas colunas são os vetores que geram o cone.)

Vamos começar por considerar o caso particular de um cone gerado por uma matriz não-singular A . Representamos por \mathbb{R}_{+0}^n o conjunto $\{x \in \mathbb{R}^n \mid x \geq 0\}$. Repare-se que este conjunto é fechado (deixado como exercício ao leitor.)

Assim sendo, queremos considerar o conjunto $A\mathbb{R}_{+0}^n$, e queremos justificar que este é fechado.

Repare-se que a função linear $x \mapsto Ax$ é contínua. Infelizmente, não é necessariamente verdade que a imagem sob uma função contínua de um conjunto fechado seja fechado. No entanto, há um resultado que nos pode ajudar:

Prop 16. *Se $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ é uma bijeção de inversa contínua e X é um conjunto fechado, $f(X)$ é fechado.*

Dem. Seja g a inversa de f , que sabemos ser contínua. Temos que $f(X) = g^{-1}(X)$. Como X é fechado, $g^{-1}(X)$ é fechado, e então $f(X)$ é fechado. \square

Repare-se que a nossa função $x \mapsto Ax$ está nestas condições, pois tem inversa contínua (em particular, linear) $x \mapsto A^{-1}x$. Assim sendo, a imagem sob A de \mathbb{R}_{+0}^n é fechada, e o nosso cone é fechado.

Este raciocínio pode ser feito, com algumas modificações, na hipótese mais fraca de A ter colunas linearmente independentes.

Prop 17. *Dizemos que um cone $A\mathbb{R}_{+0}^m$ é primitivo se as colunas de A são linearmente independentes.*

Qualquer cone primitivo é fechado.

Dem. Considere-se os vetores (colunas de A) a_1, \dots, a_m linearmente independentes, e extendenda-se este conjunto a uma base $a_1, \dots, a_m, a_{m+1}, \dots, a_n$.

Se A' é a matriz cujas colunas são a_1, \dots, a_n , o cone $A\mathbb{R}_{+0}^m$ é igual a $A'(\mathbb{R}_{+0}^m \times \{0\})$. A função $x \mapsto A'x$ tem inversa contínua, e o conjunto $\mathbb{R}_{+0}^m \times \{0\}$ é fechado, pelo que o cone original é fechado. \square

Infelizmente, não é fácil estender este raciocínio a qualquer cone. É possível, no entanto escrever qualquer cone como uma união de cones primitivos.

A ideia é a seguinte: suponhamos que tenho o cone gerado por a_1, \dots, a_m . Qualquer elemento x deste cone pode ser escrito como $t_1a_1 + \dots + t_ma_m$ para

$t_i \geq 0$, $i = 1, \dots, m$. No entanto, da Álgebra Linear, sabemos que podemos escrevê-lo como $t_1 a_{i_1} + \dots + t_k a_{i_k}$ onde a_{i_1}, \dots, a_{i_k} são linearmente independentes. Se conseguirmos assegurar que há uma forma de escrever isto de modo a que t_1, \dots, t_k são todos ≥ 0 , obtemos que x pertence ao cone (primitivo) gerado por a_{i_1}, \dots, a_{i_k} .

Feito isto, considere-se o cone gerado por $B = \{a_1, \dots, a_n\}$. O que isto provaria seria que qualquer x no cone gerado por B pertence a um cone gerado por um $B' \subseteq B$ tal que B' é linearmente independente, e então o cone gerado por B é a união dos cones gerados pelos subconjuntos linearmente independentes de B , e então é união (finita) de cones primitivos, sendo então fechado.

Falta apenas, então, mostrar que é sempre possível escrever um elemento do cone como soma não-negativa de elementos linearmente independentes.

Prop 18. *Qualquer cone é união finita de cones primitivos, sendo então fechado.*

Dem. Pelo que já vimos antes, basta mostrar que, para x da forma $t_1 a_1 + \dots + t_m a_m$, ($t_1, \dots, t_m \geq 0$), existem i_1, \dots, i_k e $r_1, \dots, r_k \geq 0$ tal que a_{i_1}, \dots, a_{i_k} são linearmente independentes e $x = r_1 a_{i_1} + \dots + r_k a_{i_k}$.

Para fazer isto, suponha-se que x pertence ao cone gerado por a_1, \dots, a_m , e considere-se uma forma de escrever x que minimize os índices diferentes de zero. Ou seja, vamos escrever x como $t_1 a_1 + \dots + t_m a_m$ de uma forma que faça com que o número de i tal que $t_i > 0$ seja mínimo.

Sejam i_1, \dots, i_k os índices tais que $t_{i_1}, \dots, t_{i_k} > 0$. Provamos que a_{i_1}, \dots, a_{i_k} são linearmente independentes.

Para este objetivo, suponha-se que não são. Mostraremos que afinal esta forma de escrever x não é o mínimo que tínhamos suposto.

Se não são linearmente independentes, existem s_1, \dots, s_k , nem todos iguais a zero (suponha-se spdg que pelo menos um deles é maior que zero), tal que $s_1 a_{i_1} + \dots + s_k a_{i_k} = 0$

Comparando as duas equações que temos até agora: (E pondo $r_j = t_{i_j}$)

$$\begin{aligned} r_1 a_{i_1} + \dots + r_k a_{i_k} &= x \\ s_1 a_{i_1} + \dots + s_k a_{i_k} &= 0 \end{aligned}$$

A ideia é subtrair a equação de baixo à de cima de modo a deixar todos os termos ≥ 0 , mas cancelar pelo menos um, e então mostrar que existe uma forma de somar os a 's a zero que tem menos termos não-nulos.

Para este efeito, considere-se o conjunto $\{\frac{r_j}{s_j} \mid s_j > 0\}$. Este conjunto é não-vazio por hipótese, e é finito, pelo que tem um mínimo $\frac{r_\ell}{s_\ell}$. Ponha-se então $r'_j = r_j - \frac{r_\ell}{s_\ell} s_j$.

É fácil verificar que $r'_j \geq 0$ para todo j (o caso em que $s_j \leq 0$ é trivial, o caso em que $s_j > 0$ é deixado ao leitor) e que $r'_\ell = 0$. Finalmente, é trivial ver que $r'_1 a_{i_1} + \dots + r'_k a_{i_k} = x$, pois o que foi feito foi subtrair a equação de baixo à equação de cima $\frac{r_\ell}{s_\ell}$ vezes, pelo que o resultado dá $x + \frac{r_\ell}{s_\ell} 0 = x$.

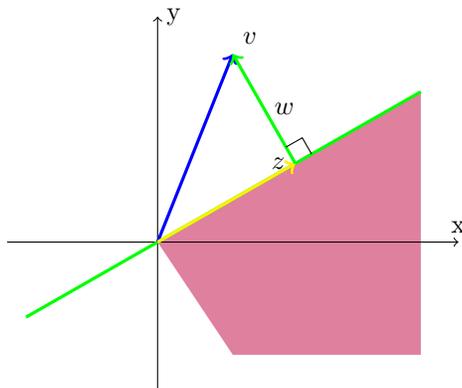
Isto mostra, então, que a forma de somar a x que minimiza o número de termos não-zero tem todos os termos linearmente independentes, pelo que x é soma de vetores linearmente independentes e então pertence a um dos cones primitivos gerados por subconjuntos linearmente independentes de $B = \{a_1, \dots, a_m\}$.

Como o número de tais subconjuntos é finito, o cone gerado por B é uma união finita de cones primitivos. Como todos estes são fechados, e uniões finitas de conjuntos fechados são fechados, o nosso cone é fechado. \square

Esta proposição é a principal razão pela qual passámos pelos cones primitivos. Visto que agora sabemos que os cones são todos fechados, sabemos que podemos considerar projeções ortogonais, e avançamos para o lema de Farkas.

6.5 O Lema de Farkas (Parte 2)

Vamos, agora, voltar aos nossos desenhos. Relembre-se que temos um ponto v e um cone C (a vermelho na imagem abaixo) e pretendemos arranjar um hiperplano que separasse v de C . Tínhamos a conjectura que o w construído abaixo (que já sabemos que podemos construir) poderia ser de uso.



Repare-se que não justificámos o ângulo reto na figura. No entanto, da figura é óbvio (e vamos demonstrar) que, de facto, $z \cdot w = 0$. Isto vai ser relevante na nossa prova.

A ideia para justificar que, para todo $x \in C$ se tem $w \cdot x \leq 0$ é a seguinte: suponhamos que havia $x \in C$ tal que $w \cdot x > 0$.

Repare-se que o cone é convexo (exercício ao leitor), pelo que todos os vetores da forma $(1-t)z+tx$ pertencem ao cone para $t \in [0, 1]$. Mas repare-se no aspeto da função $f(t) = |v - ((1-t)z+tx)|^2$. Expandido com o produto escalar, ficamos com

$$\begin{aligned} f(t) &= |(v - z) + t(z - x)|^2 \\ &= |v - z|^2 + 2(v - z) \cdot (z - x)t + |z - x|^2 t^2 \end{aligned}$$

Isto tem a aparência de uma parábola virada para cima. Agora, vejamos como é que esta se parece no intervalo $[0, 1]$... Calculando a derivada desta

função em zero, é possível verificar que esta dá $2(v - z) \cdot (z - x) = -2w \cdot x < 0$ (as contas serão feitas com mais cuidado em breve) e então a derivada é negativa em zero... O que significa que existe uma vizinhança de zero onde a função toma valores menores do que $f(0)$, o que significa que existe t_0 tal que $f(t_0) < f(0)$. Logo, temos um ponto, $(1 - t_0)z + t_0x$ que pertence ao cone, cuja distância a v é menor do que a de z , o que contraria a nossa hipótese de minimalidade quanto a z .

Isto é o cerne da demonstração. Os detalhes serão feitos agora.

Prop 19. (*Lema de Farkas*) *Seja $C \subseteq \mathbb{R}^n$ um cone, $v \in \mathbb{R}^n$ $v \notin C$ sse existe $w \in \mathbb{R}^n$ tal que $w \cdot v > 0$ e $w \cdot x \leq 0$ para todo $x \in C$.*

Dem. Primeiro, lide-se com a implicação óbvia: (\leftarrow).

Se existe w tal que $w \cdot x \leq 0$ para todo $x \in C$ e $w \cdot v > 0$, é óbvio que v não pode pertencer a C . Feito.

O que nos leva à implicação não-óbvia.

Suponhamos que $v \notin C$. Seja z a projeção ortogonal de v em C , que sabemos existir pois C é fechado e não-vazio, e seja $w = v - z$.

Para provar que $w \cdot z = 0$, considere-se a função $g(t) = |v - tz|^2$. Sabemos, por definição de z , que g tem um mínimo em $t = 1$, pois os pontos da forma tz pertencem todos ao cone para $t \in \mathbb{R}_0^+$ (verifique). Logo, $g'(1) = 0$.

Calculando $g(t)$ (expandindo o quadrado da norma com o produto escalar), temos que este dá $|v|^2 - 2(v \cdot z)t + |z|^2t^2$. A derivada é, então, $2(-(v \cdot z) + (z \cdot z)t)$. Em particular, sabemos que para $t = 1$ isto dá zero, pelo que $2(-v \cdot z + z \cdot z) = 0$ e então temos que $(v - z) \cdot z = 0$ donde se conclui $w \cdot z = 0$, como queríamos demonstrar.

Feito isto, continue-se a prova. Pretendemos mostrar que para todo $x \in C$ se tem $w \cdot x \leq 0$, pelo que supomos por contradição que existe $x \in C$ tal que $w \cdot x > 0$.

Visto que o cone é convexo (verifique) temos que para todo $t \in [0, 1]$ o vetor $(1 - t)z + tx$ pertence ao cone. Assim sendo, consideramos a função

$$\begin{aligned} f(t) &= |v - ((1 - t)z + tx)|^2 \\ &= |(v - z) + t(z - x)|^2 \\ &= |v - z|^2 + 2(v - z) \cdot (z - x)t + |z - x|^2t^2 \end{aligned}$$

Esta função tem, como já vimos, derivada negativa em zero:

$$\begin{aligned} f'(t) &= 2(v - z) \cdot (z - x) + 2|z - x|^2t \\ f'(0) &= 2(v - z) \cdot (z - x) \\ &= 2w \cdot z - 2w \cdot x \\ &= -2w \cdot x \\ &< 0 \end{aligned}$$

Logo, para $t > 0$ pequeno o suficiente tem-se $f(t) < f(0)$. Pondo t_0 um destes t pequenos o suficiente, temos que $z' = (1 - t_0)z + t_0x$ é um ponto que pertence ao cone que tem distância a v menor do que z , o que contraria a nossa hipótese sobre z . Assim sendo, a hipótese que $w \cdot x > 0$ tem que ser falsa, e o lema de Farkas está provado. \square

A seguir está uma versão equivalente do Lema de Farkas, que é mais fácil de verificar na prática, pois basta verificar a condição $w \cdot x \leq 0$ para um número finito de vetores.

Prop 20. (Lema de Farkas Modificado) *Seja A uma matriz $n \times m$, $v \in \mathbb{R}^n$ $v \notin \text{CCA}$ sse existe $w \in \mathbb{R}^n$ tal que $w \cdot v > 0$ e $A^T w \leq 0$.*

Dem. Sabemos que $v \notin \text{CCA}$ sse existe $w \in \mathbb{R}^n$ tal que $w \cdot v > 0$ e $w \cdot (Ax) \leq 0$ para todo $x \geq 0$.

Em particular, pondo $x = e_1, e_2, \dots, e_m$ (onde e_i representa o i -ésimo vetor da base canônica em \mathbb{R}^m), temos que $v \notin \text{CCA}$ implica que existe $w \in \mathbb{R}^n$ tal que $w \cdot v > 0$ e $w \cdot a_i \leq 0$ para todo $i = 1, \dots, m$. Isto pode ser reformulado como $w \cdot v > 0$ e $A^T w \leq 0$, o que conclui a implicação (\rightarrow).

Basta agora mostrar (\leftarrow). Suponha-se que existe $w \in \mathbb{R}^n$ tal que $w \cdot v > 0$ e $A^T w \leq 0$. Mostraremos que este w é tal que $w \cdot (Ax) \leq 0$ para todo $x \geq 0$, o que, por Farkas, mostra que $v \notin \text{CCA}$.

Ora, de $A^T w \leq 0$ obtemos que $w \cdot a_i \leq 0$ para todo i , onde a_i representa a i -ésima coluna de A . Assim sendo, pondo x um vetor não-negativo, temos $w \cdot (Ax) = w \cdot \sum_{i=1}^m x_i a_i = \sum_{i=1}^m x_i (w \cdot a_i)$. Todos os termos desta soma são ≤ 0 , pelo que obtemos $w \cdot (Ax) \leq 0$, como queríamos demonstrar. \square

6.6 Método das linhas ativas (Parte 2)

Recorde-se da proposição 14, que provámos anteriormente:

Prop. *Seja P o seguinte pol na forma semicanónica:*

$$\begin{cases} \max_x cx \\ Ax \leq b \end{cases}$$

E seja x um vetor admissível de P . x não é solução sse existe $w \in \mathbb{R}^n$ tal que $cw > 0$ e $A^T w \leq 0$

Com o lema de Farkas, podemos refinar isto para o seguinte útil critério:

Prop 21. (Critério das linhas ativas)

Sob as mesmas hipóteses, x é solução sse c^T pertence a CCA^{xT}

A demonstração disto é uma aplicação trivial do lema de Farkas, que o leitor poderá confirmar por si mesmo.

Visto que qualquer pol pode ser transformado na forma semicanónica, este critério acaba por ser extremamente útil para demonstrar que um dado vetor é solução.

7 Dualidade

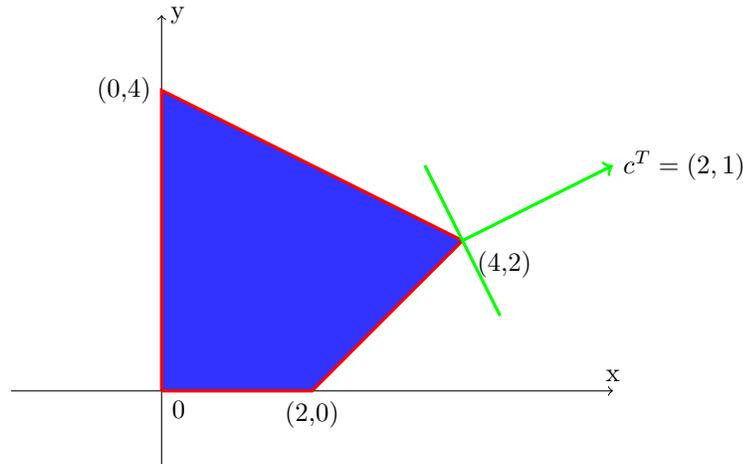
7.1 Motivação

Para motivar o conceito de dualidade, vamos primeiro considerar um exemplo.

Considere-se o seguinte pol, que já apareceu neste documento:

$$P = \begin{cases} \max_{(x,y)} 2x + y \\ x + 2y \leq 8 \\ x - y \leq 2 \\ x, y \geq 0 \end{cases}$$

Repete-se a representação visual da região admissível, do vetor c^T e da solução (4,2).



Da imagem, é claro que este ponto é solução. Mas uma imagem não é uma demonstração, e vamos então provar isto analiticamente.

Queremos provar que (4,2) é solução, ou seja, que todos os pontos em X_P têm pontuação $\leq 2 \cdot 4 + 2 = 10$.

Para fazer isto, peguemos nas desigualdades que temos: $x + 2y \leq 8$ e $x - y \leq 2$.

Normalmente, dados sistemas de equações lineares, adicionamos múltiplos de umas às outras para obter informação sobre as nossas variáveis. É isso que fazemos aqui.

Por exemplo, adicionando a primeira desigualdade mais duas vezes a segunda, obtemos $(x + 2y) + 2(x - y) \leq 8 + 2(2)$, ou seja, $3x \leq 12$, e então $x \leq 4$ para todo $(x, y) \in X_P$.

Observe-se a desigualdade $x + 2y \leq 8$, que é igual a $\frac{1}{2}x + y \leq 4$. Somando $\frac{3}{2}x$, obtemos $2x + y \leq \frac{3}{2}x + 4 \leq 6 + 4 = 10$. Logo, para todo $(x, y) \in X_P$ temos que a sua pontuação é no máximo 10, pelo que, como o nosso ponto (4,2) tem pontuação 10, este é solução.

Reiterando, o que fizemos foi pegar nas nossas desigualdades e somá-las de modo a que o lado esquerdo fosse a expressão da nossa função-objetivo, e então obter um majorante desta.

Vamos fazer outro exemplo, desta vez sem imagem:

Considere-se o pol

$$P = \begin{cases} \max_{(x,y)} x + y \\ x + 2y \leq 8 \\ 5x + y \leq 13 \\ x, y \geq 0 \end{cases}$$

Usando uma ideia semelhante, obtemos que, somando $4 \times$ a primeira desigualdade a $1 \times$ a segunda, $4(x + 2y) + 1(5x + y) \leq 4 \cdot 8 + 13$, donde $9x + 9y \leq 45$ e então $x + y \leq 5$. Agora, se arranjarmos um elemento admissível que tenha pontuação 5, sabemos que este é solução, e de facto, o ponto $(2, 3)$ é admissível, de pontuação 5, pelo que é solução.

Só mais um exemplo, este trivial para mostrar que às vezes a condição ≥ 0 é necessária. Remova-se a segunda desigualdade ao pol anterior para obter

$$P = \begin{cases} \max_{(x,y)} x + y \\ x + 2y \leq 8 \\ x, y \geq 0 \end{cases}$$

Bem, sabendo que $8 \geq x + 2y$, e sabendo que $y \geq 0$, obtemos $8 \geq x + y + y \geq x + y$, donde $x + y \leq 8$.

Assim, se arranjarmos um ponto com pontuação 8, este é solução. E de facto, pondo $(x, y) = (8, 0)$, temos a nossa solução desejada.

Isto já são exemplos suficientes para fazer o seguinte ponto: adicionar desigualdades é útil. E é nisso que se baseia a dualidade.

7.2 Introdução à dualidade

Considere-se um pol P na forma canónica:

$$\begin{cases} \max_x cx \\ Ax \leq b \\ x \geq 0 \end{cases}$$

Vamos tentar generalizar o método de somar desigualdades. Vamos escrever a desigualdade $Ax \leq b$ como:

$$\begin{aligned}
a_1x &\leq b_1 \\
a_2x &\leq b_2 \\
&\dots \\
a_mx &\leq b_m
\end{aligned}$$

Sejam y_1, \dots, y_m escalares ≥ 0 (isto é necessário porque multiplicar por escalares menores que zero troca as desigualdades). Temos que

$$y_1(a_1x) + y_2(a_2x) + \dots + y_m(a_mx) \leq y_1b_1 + y_2b_2 + \dots + y_mb_m$$

Isto pode ser escrito de forma compacta da seguinte forma. Seja y o vetor -coluna cuja i -ésima componente é y_i . Temos que, se $y \geq 0$,

$$y^T Ax \leq y^T b$$

Visto que isto é um número real, transpôr não faz diferença, pelo que

$$b^T y \geq y^T Ax$$

Relembre-se que esta mixórdia de somar desigualdades umas às outras foi com o intuito de arranjar majorantes de cx . Portanto, sendo p_i a i -ésima casa de $y^T A$, isto dá-nos que

$$b^T y \geq p_1x_1 + p_2x_2 + \dots + p_nx_n$$

Aqui podemos usar o facto que $x \geq 0$ para obter que, *na condição que $p_i \geq c_i$ para todo i ,*

$$\begin{aligned}
b^T y &\geq p_1x_1 + p_2x_2 + \dots + p_nx_n \\
&\geq c_1x_1 + c_2x_2 + \dots + c_nx_n = cx
\end{aligned}$$

E então $b^T y$ é um majorante da função objetivo.

Em suma: dado o pol

$$P = \begin{cases} \max_x cx \\ Ax \leq b \\ x \geq 0 \end{cases}$$

É possível arranjar majorantes de cx somando as desigualdades de $Ax \leq b$. Na condição de $A^T y \geq c^T$, temos que $b^T y$ é majorante de cx .

Obviamente estamos interessados no *menor* majorante, daí fazer sentido considerarmos o pol

$$\begin{cases} \min_y b^T y \\ A^T y \geq c^T \\ y \geq 0 \end{cases}$$

Este é o chamado *dual de P*.

Um facto interessante: se a mesma estratégia for aplicada a este pol, acabamos com o pol com que começámos!

Nomeadamente, considere-se somar as linhas de $A^T y \geq c^T$ com o vetor $z \geq 0$. Ficamos com $z^T A^T y \geq z^T c^T = cz$, pelo que, na condição de $z^T A^T$ ter todas as respetivas componentes menores ou iguais do que b^T , cz é minorante de $b^T y$. Ou seja, para obter o maior minorante, resolvemos o pol

$$\begin{cases} \max_z cz \\ Az \leq b \\ z \geq 0 \end{cases}$$

Que é exatamente o pol com que começámos.

7.3 Algumas palavras sobre o dual

Antes de avançar na teoria, gostaria de fazer notar uma coisa.

A técnica usada anteriormente, de somar desigualdades, é útil, mas não é um método rigoroso de descobrir ‘o dual de um pol’. Até agora foi considerado o pol na forma canónica, mas vamos adiante considerar duais de pols na forma padrão e semicanónica. E um facto importante é que o dual do mesmo pol em qualquer uma destas três formas é diferente. Ou seja, o dual depende da forma como o pol é expresso. E, por exemplo, se tentássemos fazer o dual do pol

$$\begin{cases} \min_{(x,y)} -x + y \\ -2x - y \geq 0 \\ x \geq 0 \\ y \geq 0 \end{cases}$$

Teríamos resultados diferentes se interpretássemos este como um pol da forma

$$\begin{cases} \min_y cy \\ Ay \geq b \\ y \geq 0 \end{cases}$$

Ou da forma

$$\begin{cases} \min_y cy \\ Ay \geq b \end{cases}$$

Portanto o dual não depende só da forma como o pol está expresso, mas também da forma como o interpretamos... Assim sendo, reservamos o termo ‘o dual de um pol’ como um termo informal, cujo significado será óbvio de contexto e definições anteriores, e não como uma definição rigorosa.

Para terminar esta parte, o termo dual tem significado rigoroso no contexto de problemas de max-min e min-max. Ou seja, problemas da forma:

$$\sup_x \inf_y \mathcal{L}(x, y) \text{ ou } \inf_y \sup_x \mathcal{L}(x, y)$$

Em que \mathcal{L} é uma função $X \times Y \rightarrow \mathbb{R}$ e X e Y são conjuntos arbitrários.

Neste contexto, o dual é não-ambíguo: estes dois problemas são duais um do outro.

É possível tentar fazer a noção de dual no contexto do pol mais rigorosa, transformando o pol num problema max-min e min-max. No entanto, esta transformação sofre dos mesmos problemas de depender de como interpretamos o pol, e é bastante artificial.

7.4 Propriedades básicas do dual

Relembre-se que dado o pol na forma canónica

$$P = \begin{cases} \max_x cx \\ Ax \leq b \\ x \geq 0 \end{cases}$$

Definimos o seu dual como

$$Q = \begin{cases} \min_y b^T y \\ A^T y \geq c^T \\ y \geq 0 \end{cases}$$

Para simplificar a notação, vamos, de agora em diante, usar X e S para nos referirmos ao conjunto admissível e solução do pol em questão, e Y e R para os respetivos conjuntos do problema dual.

Primeiro que tudo, a proposição que motivou a consideração do problema dual.

Prop 22. (*Teorema fraco da dualidade*)

Se $x \in X$ e $y \in Y$, então $cx \leq b^T y$.

Como consequência, se $cx = b^T y$, x e y são soluções dos respetivos problemas.

Dem. Se x é admissível, $Ax \leq b$, donde, por y ser ≥ 0 , temos $y^T Ax \leq y^T b$. Devido a estes serem números reais, podemos transpô-los, para obter $x^T A^T y \leq b^T y$. Como $c^T \leq A^T y$ e $x \geq 0$, temos $x^T c^T \leq x^T A^T y \leq b^T y$ e então, transpondo, obtemos $cx \leq b^T y$, como pretendíamos demonstrar.

A demonstração de que se $cx = b^T y$ então estes são soluções é trivial e deixada como exercício ao leitor. \square

O pol dual pode também ser utilizado para servir de prova de existência de solução.

Prop 23. *Se X e Y forem ambos não-vazios, tanto o pol original como o seu dual têm solução.*

Dem. Recorde-se que, visto que o pol original é um pol de maximização, se a função-objetivo for majorada, tem solução. Se Y é não-vazio, qualquer seu elemento tem pontuação que majora a dos elementos de X , pelo que o pol original tem pontuação majorada e então tem solução.

A prova que o dual tem solução é semelhante e deixada para o leitor. \square

Recorde-se que nos exemplos que fizemos, era sempre possível somar as desigualdades de modo a obter um majorante minimal. Ou seja, nesses exemplos, conseguimos sempre descobrir um upper bound que corresponde à solução.

Não é inconcebível de um cenário em que isto é impossível. Ou seja, um pol em que a solução é, digamos, 2, mas com a estratégia do dual só conseguimos provar que é no máximo 3.

Acontece que isto é, de facto, impossível. Ou seja, tendo como base a estratégia do dual, o menor majorante corresponde exatamente à solução.

Para provar isto, provamos primeiro uma coisa parecida: caso exista solução, é possível arranjar majorantes tão ‘apertados’ quanto queiramos.

Prop 24. *Seja s uma solução do pol original. Então, para todo $\varepsilon > 0$ existe $y \in Y$ tal que $cs \leq b^T y < cs + \varepsilon$*

Da mesma forma, se r é solução do dual, para todo ε existe $x \in X$ tal que $b^T y - \varepsilon < cx \leq b^T y$.

Dem. Suponha-se que o pol original é

$$P = \begin{cases} \max_x cx \\ Ax \leq b \\ x \geq 0 \end{cases}$$

E que este tem uma solução s .

A ideia para criar um majorante arbitrariamente próximo é usar o lema de Farkas de forma inteligente.

Repare-se que não existe x tal que

$$\begin{aligned} Ax &\leq b \\ x &\geq 0 \\ cx &\geq cs + \varepsilon \end{aligned}$$

Repare-se que isto é a mesma coisa que dizer que não existe $x \geq 0$ tal que

$$\begin{bmatrix} A \\ -c \end{bmatrix} x \leq \begin{bmatrix} b \\ -cs - \varepsilon \end{bmatrix}$$

Isto pode ser escrito como uma afirmação sobre cones: adicionando as variáveis de folga f , temos que isto é igual a dizer que não existe $x, f \geq 0$ tal que

$$\begin{bmatrix} A & I \\ -c & 1 \end{bmatrix} \begin{bmatrix} x \\ f \end{bmatrix} = \begin{bmatrix} b \\ -cs - \varepsilon \end{bmatrix}$$

Isto é o mesmo que dizer que o vetor $(b, -cs - \varepsilon)$ não pertence a $\text{CC} \begin{bmatrix} A & I \\ -c & 1 \end{bmatrix}$.

Isto permite-nos aplicar Farkas e dizer que existe um vetor (w, z) tal que

$$(w, z) \cdot (b, -cs - \varepsilon) < 0$$

E

$$\begin{bmatrix} A^T & -c^T \\ I & 1 \end{bmatrix} (w, z) \geq 0$$

Repare-se que trocámos as desigualdades do lema de Farkas. Isto foi por uma razão de conveniência, pelo que se vai ver no que se segue. Não há problema em fazer isto, pois basta pegar no vetor assegurado pelo lema de Farkas e considerar a sua negação.

Rearranjando os termos acima usando multiplicação de matrizes por blocos, obtemos que (w, z) satisfaz:

$$\begin{aligned} b^T w &< z(cx + \varepsilon) \\ A^T w &\geq zc^T \\ w &\geq 0 \\ z &\geq 0 \end{aligned}$$

É fácil, agora, verificar que, se por acaso z for diferente de zero, o vetor $y = \frac{1}{z}w$ satisfaz o que é dele desejado. Basta, então, mostrar que $z \neq 0$.

Para ver isto, repare-se que, se z fosse zero, ter-se-ia que

$$\begin{aligned} b^T w &< 0 \\ A^T w &\geq c^T \\ w &\geq 0 \end{aligned}$$

O que, usando o lema de Farkas, implicaria que $b \notin \text{CC} [A \ I]$. Usando o truque das variáveis de folga, é fácil ver que isto é o mesmo que dizer que não

existe $x \geq 0$ tal que $Ax \leq b$, mas sabemos que isto é falso, pois por hipótese, o nosso vetor s é admissível. Isto mostra que z não pode ser igual a zero, e então existe o nosso vetor y desejado.

A prova da segunda parte pode ser feita de forma análoga, ou pode ser feita reparando que o pol dual é equivalente a

$$\begin{cases} \max_y -b^T y \\ -A^T y \leq -c^T \\ y \geq 0 \end{cases}$$

E que o dual deste (interpretando este como um pol na forma canónica) é equivalente ao pol original.

Então, usando esta proposição, sendo r uma solução do dual, temos que existe x admissível no pol original tal que $-b^T y \leq -cx < -b^T y + \varepsilon$, ou seja, existe x admissível tal que $b^T y - \varepsilon < cx \leq b^T y$, como pretendíamos demonstrar. \square

Isto leva diretamente a:

Prop 25. (Teorema forte da dualidade)

Fixo o pol na forma canónica e o seu dual, um dos seguintes quatro casos acontece necessariamente:

- i* $X = \emptyset = Y$ e então nenhum dos pols tem solução.
- ii* $X = \emptyset \neq Y$ e então a função-objetivo não é minorada no pol dual e nenhum deles tem solução.
- iii* $X \neq \emptyset = Y$ e então a função-objetivo não é majorada no pol original e nenhum deles tem solução.
- iv* $X \neq \emptyset \neq Y$ e então ambos os pols têm solução. Se $s \in S$ e $r \in R$, temos $cs = b^T r$.

Dem. Esta prova vai examinar os quatro casos separadamente.

i) Este caso é trivial.

ii) Sabendo que $Y \neq \emptyset$, se a função-objetivo fosse minorada no dual ter-se-ia que este tem solução. Pela prop anterior, isto implicaria $X \neq \emptyset$.

iii) Análogo ao anterior.

iv) Suponha-se que X e Y são não-vazios. Então pelo teorema fraco da dualidade, ambos os pols têm solução, digamos s e r .

Temos que $cs = b^T r$ pois, caso contrário, seja $\varepsilon = b^T r - cs$. Sabemos pela prop anterior que existe y tal que $b^T y < cs + \varepsilon = b^T r$, o que contraria a hipótese de r ser solução. Logo, $cs = b^T r$. \square

7.5 Outros duais

Até agora, já tivemos bastante sucesso a arranjar formas de verificar se um vetor é solução de um pol: o método das linhas ativas para pols na forma semi-canónica e dualidade para pols na forma canónica. Isto ser-nos-ia útil se tivessemos uma forma de *adivinhar* soluções. Temos um método (vetores básicos), mas este só funciona para o pol na forma padrão. Assim sendo, faz sentido tentar estender os nossos métodos de verificação ao pol na forma padrão.

Vamos investigar dualidade para o pol na forma padrão. Claro que é sempre possível transformar o pol na forma padrão para o pol na forma canónica e tomar o dual deste – E vamos investigar isto na sequência – mas vamos tentar fazer primeiro as coisas de raiz como fizemos com o canónico.

Para exemplificar e depois generalizar, vamos considerar um pol na forma padrão:

$$\begin{cases} \min_{(x,y,z)} x + 2y - z \\ x + y + z = 1 \\ x, y, z \geq 0 \end{cases}$$

Da primeira restrição, sabemos que $x + y + z = 1$ e então $-x - y - z = -1$. Ora, $x + 2y - z \geq -x - y - z = -1$, donde a função-objetivo é sempre maior ou igual que -1 .

Como o ponto $(0, 0, 1)$ toma este valor, sabemos que este é solução.

Para um exemplo menos trivial, vamos adicionar uma restrição a este problema.

$$\begin{cases} \min_{(x,y,z)} x + 2y - z \\ x + y + z = 1 \\ -y + z = 0 \\ x, y, z \geq 0 \end{cases}$$

Desta vez, repare-se que $x + 2y - z \geq \frac{1}{2}x + 2y - z = \frac{1}{2}(x + y + z) - \frac{3}{2}(-y + z) = \frac{1}{2}$. Logo, se arranjarmos um elemento que tenha pontuação $\frac{1}{2}$ este tem que ser solução, e o vetor $(0, \frac{1}{2}, \frac{1}{2})$ é admissível e tem esta pontuação, sendo então solução.

Vamos agora generalizar. Suponhamos que temos um pol na forma padrão:

$$\begin{cases} \min_x cx \\ Ax = b \\ x \geq 0 \end{cases}$$

Somando as linhas de $Ax = b$, obtemos igualdades da forma $y^T Ax = y^T b$. Ao contrario do pol na forma canónica, não nos precisamos aqui de preocupar com que y seja ≥ 0 , pois não há sinais de desigualdades a trocar.

Queremos arranjar minorantes da função objetivo, portanto queremos que $y^T Ax \leq cx$. Para isso, é suficiente que $y^T A \leq c$, e então temos o nosso minorante: $b^T y = y^T Ax \leq cx$ desde que $A^T y \leq c^T$

Claro que queremos arranjar o maior minorante possível, pelo que ficamos que o dual do pol na forma padrão:

$$\begin{cases} \max_y b^T y \\ A^T y \leq c^T \end{cases}$$

O leitor poderá reparar que isto é um pol na forma semicanónica.

Fazemos a mesma estratégia na direção oposta, para justificar que o dual do pol na forma semicanónica é o pol original na forma padrão.

Podemos somar linhas da desigualdade $A^T y \leq c^T$ para obter $x^T A^T y \leq x^T c^T$. Se $Ax = b$, temos que $cx \geq (Ax)^T y = b^T y$, donde cx é majorante de $b^T y$. Queremos o menor majorante possível, pelo que voltamos a

$$\begin{cases} \min_x cx \\ Ax = b \\ x \geq 0 \end{cases}$$

Como desejado.

Queremos agora justificar que estes dois pols têm propriedades que já associámos ao dual. Nomeadamente, queremos demonstrar algo parecido com o teorema forte da dualidade.

Podemos fazer isto diretamente, claro, mas podemos ser mais económicos aproveitando o teorema forte da dualidade que já temos.

Vamos fazer isto usando traduções: vamos traduzir ambos os pols para pols de formas em que podemos dizer que estes são o dual um do outro.

Primeiro que tudo, repare-se que o pol

$$P = \begin{cases} \min_x cx \\ Ax = b \\ x \geq 0 \end{cases}$$

É equivalente ao pol na forma canónica

$$P' = \begin{cases} \max_x -cx \\ \begin{bmatrix} -A \\ A \end{bmatrix} x \leq \begin{bmatrix} -b \\ b \end{bmatrix} \\ x \geq 0 \end{cases}$$

Que por sua vez tem dual

$$Q' = \begin{cases} \min_{(y_+, y_-)} -b^T(y_+ - y_-) \\ -A^T(y_+ - y_-) \geq -c^T \\ y_+, y_- \geq 0 \end{cases}$$

Mas este pol é uma tradução do pol

$$Q = \begin{cases} \max_y b^T y \\ A^T y \leq c^T \end{cases}$$

Como descrito na secção Tradução - Adição de positividades.

Assim sendo, temos que $X_P = \emptyset$ sse $X_{P'} = \emptyset$, $S_P = \emptyset$ sse $S_{P'} = \emptyset$, e coisas semelhantes para Q e Q' . Assim sendo, aplicando o teorema forte da dualidade a P' e Q' , temos que um se algum de X_P , X_Q for vazio, nenhum dos quatro pols acima tem solução. Se X_P e X_Q forem não-vazios, $X_{P'}$ e $X_{Q'}$ são não-vazios e então, por dualidade, $S_{P'}$ e $S_{Q'}$ são não-vazios, e por tradução temos que P e Q têm ambos solução.

Sabendo que o valor mínimo de P é a negação do valor máximo de P' , que por sua vez é igual ao valor mínimo de Q' que é a negação do valor máximo de Q . Logo, se $s \in S_P$ e $r \in S_Q$ temos que $cs = b^T r$

Isto acaba o nosso esboço de demonstração do teorema forte da dualidade para o pol na forma padrão. O leitor é, no entanto, encorajado a verificar os detalhes, e/ou, se tiver pachorra (o autor não teve), a fazer a demonstração com o lema de Farkas.

Resumindo isto tudo:

Prop 26. (Teorema forte da dualidade para o pol na forma padrão)

Seja P o pol

$$\begin{cases} \min_x cx \\ Ax = b \\ x \geq 0 \end{cases}$$

E Q o pol

$$\begin{cases} \max_y b^T y \\ A^T y \leq c^T \end{cases}$$

Dizemos que estes são duais um do outro.

Se algum destes não tiver vetores admissíveis, nenhum dos dois tem solução. Se ambos tiverem vetores admissíveis, têm ambos solução e se s é solução de P e r é solução de Q tem-se $cs = b^T r$.

8 Simplexo

8.1 Introdução

O algoritmo do Simplexo é bastante importante, visto que permite resolver problemas de otimização linear muito mais rapidamente do que os métodos que o precederam. E apesar de haver outros algoritmos, na teoria mais rápidos do que o Simplexo², este acaba por ser usado ainda hoje em aplicações práticas, dado que é, na vasta maioria dos casos, extremamente eficiente.

A ideia base por trás do Simplexo é razoavelmente simples. Considere-se o algoritmo dos vetores básicos para o pol na forma padrão. Visto que sabemos que, se houver solução, existe solução básica, basta verificar os vetores básicos todos, e destes o que tiver melhor pontuação é o nosso candidato a solução.

Este algoritmo tem dois problemas: primeiro que tudo, não nos diz se um dado pol tem de facto solução, apenas indica o melhor candidato para tal. Segundo, um pol pode ter imensos vetores admissíveis básicos, pelo que pode ser muito ineficiente verificá-los a todos.

Veremos que, na tentativa de resolver um destes problemas, acabamos por resolver o outro, e o resultado final será o algoritmo do Simplexo.

8.2 Motivação

A nossa procura começa com o seguinte: tentaremos aplicar dualidade para obter um teste que nos diga se, dado um pol e um vetor admissível x , se x é solução.

Recorde-se que já temos o método das linhas ativas, mas esse tem um problema: requer uma ‘testemunha’. Ou seja, ele permite-me justificar que x é solução, mas para isso preciso de mostrar que c^T pertence ao cone das colunas de A^{xT} . Para isso, preciso de arranjar um $w \geq 0$ tal que $A^{xT}w = c^T$. Arranjar este w não é fácil em geral.

Vamos, então, procurar um método que nos permita, dado um x , sob certas condições, determinar se x é solução ou não.

Consideremos então o pol na forma padrão:

$$\begin{cases} \min_x cx \\ Ax = b \\ x \geq 0 \end{cases}$$

Tendo em conta o teorema forte da dualidade, pretendemos descobrir se existe y tal que $A^T y \leq c^T$ e $b^T y = cx$. Tendo em conta que o nosso candidato

²Os detalhes serão especificados quando se falar sobre complexidade de algoritmos, mas em termos técnicos, o algoritmo do Simplexo é *worst-case* exponencial. Existem algoritmos que resolvem polys em *worst-case* polinomial. No entanto, repare-se que os casos em que o Simplexo corre lentamente são raros, de modo a que, em média (para diversos sentidos da palavra), o Simplexo é polinomial. Estes assuntos estão, no entanto, bastante fora do âmbito da cadeira, pelo que ficam apenas aqui como trivias para o leitor.

a solução x é admissível, temos $b^T = x^T A^T$, donde temos que pretendemos descobrir y tal que $x^T A^T y = x^T c^T$ e $A^T y \leq c^T$.

Uma possível ideia seria tentar descobrir y tal que $A^T y = c^T$. Tal y certamente obedeceria a ambas as condições. Infelizmente, em geral, este y não existe, visto que a matriz A^T pode ser muito alta, e então a equação $A^T y = c^T$ pode não ter solução.

Suponhamos, no entanto, que x é básico. Então, $cx = x_{P_x}^T c_{P_x}^T$. Da mesma forma, $b^T y = x^T A^T y = x_{P_x}^T (A^T y)_{P_x}$, e este último, como o leitor poderá facilmente verificar, é igual a $x_{P_x}^T A_{P_x}^T y$.³ Assim sendo, a condição $b^T y = cx$ acaba por ser equivalente a $x_{P_x}^T A_{P_x}^T y = x_{P_x}^T c_{P_x}^T$. Ficamos então com a ideia de resolver $A_{P_x}^T y = c_{P_x}^T$, e esta equação sim é solúvel, pois, por definição de vetor básico, as linhas de $A_{P_x}^T$ são linearmente independentes, e então a característica desta matriz é igual à sua altura.

Infelizmente, resolver esta equação não nos assegura nada sobre se x é solução ou não, visto que também é preciso verificar $A^T y \leq c^T$. No entanto, se se encontrar uma solução de $A_{P_x}^T y = c_{P_x}^T$ que satisfaz $A^T y \leq c^T$, de certeza que x é solução.

A proposição seguinte diz-nos que não só a existência de tal y é suficiente para que x seja solução, é também necessária.

Prop 27. *Dado o pol*

$$\begin{cases} \min_x cx \\ Ax = b \\ x \geq 0 \end{cases}$$

É um vetor x admissível,⁴ temos que x é solução sse existe y tal que $A_{P_x}^T y = c_{P_x}^T$ e $A^T y \leq c^T$.

Dem. A implicação (\leftarrow) é trivial, visto que um y tal que $A^T y \leq c^T$ é admissível do dual, e se $A_{P_x}^T y = c_{P_x}^T$ temos que $x_{P_x}^T A_{P_x}^T y = x_{P_x}^T c_{P_x}^T$ e então (verifique) $x^T A^T y = cx$ e então, visto que x é admissível, temos que $b^T y = cx$, o que, por dualidade, nos indica que x é solução.

Para a implicação (\rightarrow), aplique-se o teorema da dualidade forte. Sabemos que se x é solução existe y tal que $A^T y \leq c^T$ e $b^T y = cx$. Mas é fácil ver que esta última condição é equivalente a $x_{P_x}^T A_{P_x}^T y = x_{P_x}^T c_{P_x}^T$.

Mostraremos agora que $A_{P_x}^T y = c_{P_x}^T$. Sabemos já que $A_{P_x}^T y \leq c_{P_x}^T$, por y ser admissível do dual, e vejamos então o que acontece se não houver igualdade.

Visto que $x^T A^T y = x^T c^T$, temos que

$$\sum_{i=1}^n x_i (A^T y)_i = \sum_{i=0}^n x_i c_i$$

³Para evitar ambiguidade: na expressão $A_{P_x}^T$, o subscrito toma precedência. Ou seja, por $A_{P_x}^T$ entende-se: a matriz obtida por transposição a partir da matriz A_{P_x} . Por outras palavras, a matriz obtida de A^T apagando as linhas i tal que $x_i = 0$.

⁴Repare-se que não é tomado como hipótese que x seja básico.

Supondo que não se tem $A_{P_x}^T y = c_{P_x}^T$, existirá $j \in P_x$ tal que $(A^T y)_j \neq c_j$, e então $(A^T y)_j < c_j$. Separando este termo do resto, temos

$$x_j(A^T y)_j + \sum_{i \neq j} x_i(A^T y)_i = x_j c_j + \sum_{i \neq j} x_i c_i$$

Mas visto que todos os x_i são maiores ou iguais que zero, e como $A^T y \leq c^T$, temos que $x_i(A^T y)_i \leq x_i c_i$ para todo i , e então

$$\sum_{i \neq j} x_i(A^T y)_i \leq \sum_{i \neq j} x_i c_i$$

E como x_j é positivo e, por hipótese, $(A^T y)_j < c_j$, temos que $x_j(A^T y)_j < x_j c_j$. Somando estas duas desigualdades, obtemos

$$x_j(A^T y)_j + \sum_{i \neq j} x_i(A^T y)_i < x_j c_j + \sum_{i \neq j} x_i c_i$$

O que contradiz a igualdade acima. Logo, não pode existir tal j , donde obtemos que $A_{P_x}^T y = c_{P_x}^T$, como pretendíamos demonstrar. \square

Este novo critério para verificar se um dado x é solução é, em geral, não muito mais útil do que os que temos até agora. Existe um caso, no entanto, em que funciona na perfeição.

Suponha-se que x é um vetor admissível básico. Então, decerto que a equação $A_{P_x}^T y = c_{P_x}^T$ é possível. Verificar se existe uma solução que obedece a $A^T y \leq c^T$ é, em geral, difícil, mas se houver uma única solução, fica uma tarefa trivial. Ora, isto acontece precisamente quando $A_{P_x}^T$ é quadrada, ou seja, $\#P_x = m$, onde m é o número de linhas da matriz A .

Se x é um vetor básico tal que $\#P_x = m$, dizemos que x é um *vetor básico não-degenerado*. O que vimos agora é um teste rápido e eficaz de testar se um vetor básico não-degenerado é solução:

Resolver a equação $A_{P_x}^T y = c_{P_x}^T$

Verificar se $A^T y \leq c^T$

Caso afirmativo, x é solução

Caso contrário, x não é solução

A restrição que x seja não-degenerado não é tão restritiva quanto possa parecer. Em geral, vetores degenerados são raros, especialmente quando se trabalha com aproximações e dados reais.

Para ter uma ideia de porquê (mas repare-se que o que se segue não é uma prova formal) vamos considerar o menor caso não-trivial de um vetor básico degenerado: considere-se um pol em \mathbb{R}^3 cuja matriz A seja uma matriz 2×3 de linhas linearmente independentes. O conjunto-admissível será então uma reta.

Neste contexto, um vetor básico degenerado seria um vetor com 0 ou 1 coordenadas iguais a zero. Ora, a existência de tal vetor requer que a que a reta que representa o conjunto admissível passe por um dos eixos. O leitor pode facilmente convencer-se que, na prática, é preciso muita pontaria para que a

reta toque nos eixos. Acontece (mas não será justificado) que isto é verdade em geral, ou seja, é muito raro haver vetores básicos degenerados. Para mais, mesmo que haja, a mais pequena perturbância (e.g. um erro de arredondamento) pode transformar o problema de volta num problema não-degenerado. Assim sendo, vamos desenvolver o algoritmo do Simplexo ignorando a sua existência. É possível adaptar o algoritmo para contextos em que vetores degenerados aparecem, mas isto encontra-se fora do âmbito da cadeira.

Vamos, então, trabalhar sob a seguinte hipótese: dizemos que um pol P é *não-degenerado no sentido lato* se não tiver nenhum vetor básico degenerado.

Neste contexto, o teste acima, para além de eficaz e eficiente, dá-nos uma forma de, em caso negativo, arranjar um vetor melhor, como se verá no seguimento. É, aliás, nisto que se baseia o algoritmo do Simplexo. Os detalhes serão feitos a seguir, mas a ideia é um ciclo da forma:

```

Começar com um vetor básico  $x$ 
 $x$  é solução?
Se sim, parar.
Se não, arranjar um vetor melhor.
  Conseguimos arranjar vetores arbitrariamente bons?
  Se sim, parar. Não há solução.
  Se não, (provaremos) é possível arranjar um vetor
    básico  $v$  de pontuação melhor do que  $x$ .
    Voltar ao início, com  $x = v$ 

```

8.3 Concretização

Para finalizar o algoritmo, falta-nos apenas um ingrediente: se soubermos que x não é solução, como arranjar um vetor de melhor pontuação?

Por outras palavras, o que desejamos é arranjar um vetor \tilde{x} tal que \tilde{x} seja também admissível e tal que $c\tilde{x} < cx$.

Para arranjar um candidato, voltemos à nossa forma de justificar que um vetor x não é solução. Supondo que x é um vetor básico não-degenerado, resolvemos a equação $A_{P_x}^T y = c_{P_x}^T$ e verificámos que $A^T y \not\leq c^T$. Assim sendo, existe uma casa, digamos a j -ésima, tal que $(A^T y)_j > c_j$. Sabemos decerto que $j \notin P_x$, pois para estes índices há igualdade.

Considere-se o que acontece quando tentamos ‘andar um pouco na j -ésima direção’. Seja e_j o vetor unitário com 1 na casa j . Queremos que o nosso h seja ‘parecido com’ algo da forma te_j , mas infelizmente o vetor $x + h$ assim obtido pode não ser admissível.

Para o ser, é preciso o requerimento que $Ah = 0$ (verifique), portanto tente-se modificar o vetor e_j ligeiramente para isto acontecer. Considere-se o vetor \hat{h} tal que $\hat{h}_i = 0$ para $i \notin P_x$ e tal que $A_{P_x} \hat{h}_{P_x} = Ae_j$. Então, o vetor $h = \hat{h} - e_j$ satisfaz $Ah = 0$. Para mais, se se considerar $\tilde{x} = x - \varepsilon h$ para $\varepsilon > 0$ pequeno o suficiente, $\tilde{x} \geq 0$. Para ver porquê, repare-se que todas as casas i de \tilde{x} obedecem a um dos três casos seguintes:

1. Se $i \notin P_x$ e $i \neq j$, $\tilde{x}_i = x_i = 0 \geq 0$
2. Se $i = j$, $\tilde{x}_i = x_i + \varepsilon = \varepsilon \geq 0$
3. Se $i \in P_x$, $\tilde{x}_i = x_i - \varepsilon h_i$, e como $x_i > 0$, para ε pequeno o suficiente isto continua a ser ≥ 0 .

Um detalhe importante: se $h_i \geq 0$ para todo $i \in P_x$, *qualquer* ε funciona. No caso contrário, existe um ε maximal, que será útil na sequência.

Sabendo que \tilde{x} é admissível, basta agora mostrar que a sua pontuação é melhor do que a de x . Para isso, basta mostrar que $ch > 0$. Para esse efeito, eis as seguintes contas:

$$\begin{aligned}
 ch &= \hat{c}h - c_j \\
 &= c_{P_x} \hat{h}_{P_x} - c_j \\
 &= y^T A_{P_x} \hat{h}_{P_x} - c_j \\
 &= y^T A e_j - c_j \\
 &= (A^T y)_j - c_j \\
 &> 0
 \end{aligned}$$

O que, agradavelmente, é precisamente o desejado. Assim sendo, temos uma forma de arranjar vetores admissíveis melhores que x .

Para terminar rapidamente um caso, repare-se no seguinte: se $h \leq 0$, $x - \varepsilon h$ é admissível *para qualquer* ε positivo. Assim sendo, é possível arranjar vetores arbitrariamente bons: basta fazer $\varepsilon \rightarrow \infty$, obtendo-se que $x - \varepsilon h$ é sempre admissível, e a sua pontuação é $cx - \varepsilon ch \rightarrow -\infty$. Disto se conclui o seguinte:

Se o vetor h obtido anteriormente for tal que $h \leq 0$, o pol não tem solução.

Basta agora terminar o outro caso: supondo que $h \not\leq 0$, como arranjar um vetor melhor? De preferência, um vetor básico não-degenerado, de modo a podermos repetir o algoritmo.

Recorde-se que a razão pela qual, neste caso, não se pode tomar valores de ε arbitrariamente grandes é porque para ε grande o suficiente o vetor $x - \varepsilon h$ pode não ser ≥ 0 .

Suponha-se que i é tal que $h_i > 0$. Então, sendo $\xi_i = \frac{x_i}{h_i}$, o sinal de $x_i - \varepsilon h_i$ depende de se ε é menor, maior, ou igual a ξ_i .

O maior ε tal que $x_i - \varepsilon h_i \geq 0$ é precisamente ξ_i . Considere-se, então, $\xi = \min_{i \in P_h} \xi_i$, suponhamos ξ_k . Então, é fácil ver que, para $\varepsilon = \xi$ se tem $x - \varepsilon h \geq 0$.

Para ver isto, considere-se a divisão em casos feita acima. Para $i \notin P_x$ é trivialmente verdade que $x_i - \varepsilon h_i \geq 0$; para $i \in P_x$ isto é verdade desde que $h_i \leq 0$, e finalmente, no caso em que $h_i > 0$ temos que

$$x_i - \varepsilon h_i = x_i - \xi h_i \geq x_i - \xi_i h_i = 0$$

Em particular, para $i = k$, $x_i - \varepsilon h_i = 0$.

Assim sendo, repare-se que não só é o vetor obtido admissível, mas também é tal que as suas coordenadas positivas são as de x , com a adição do índice j e com a exclusão do índice k e possivelmente mais. Assim sendo, o número de coordenadas positivas não aumentou, e há esperanças que o vetor resultante seja básico e não-degenerado. É este o nosso candidato para ‘sucessor’.

Isto deixa-nos, no entanto, com uma primeira versão do algoritmo do simplexo:

Começar com um vetor básico não-degenerado x

Seja y a solução de $A_{P_x}^T y = c_{P_x}^T$

Se $A^T y \leq c^T$:

x é solução. Parar.

Caso contrário:

Seja j um índice tal que $(A^T y)_j > c_j$

Seja h o vetor definido da seguinte forma:

$$h_j = -1$$

$$h_{P_x} \text{ é a solução de } A_{P_x} h_{P_x} = A e_j$$

$$h_i = 0 \text{ para o resto dos índices}$$

Se $h \leq 0$:

Há vetores admissíveis arbitrariamente bons, pelo que o pol não tem solução. Parar.

Caso contrário:

Seja ξ o menor número da forma x_i/h_i , para $i \in P_h$

Seja $\tilde{x} = x - \xi h$

Se \tilde{x} é básico não-degenerado:

Começar o algoritmo de novo com $x = \tilde{x}$

Caso contrário:

Desistir. :(

Seria agora desejável provar que desistências não acontecem. Infelizmente, em geral, isto não acontece. Provaremos na secção seguinte que \tilde{x} é sempre básico, mas nada o proíbe, em geral, de ser degenerado.

Claro que se, por acaso, o nosso pol for não-degenerado, isso nunca acontecerá. E como confiamos que na prática pols degenerados quase nunca acontecem, ficamos satisfeitos com o nosso algoritmo.

8.4 Pontas soltas

Nesta secção provaremos que, se o nosso pol for não-degenerado, o algoritmo descrito acima termina sempre e nunca nos tomamos por desistentes. Ilustramos também um critério ligeiramente parvo, mas frequentemente útil em testes da Sernadas, para justificar que um pol é não-degenerado.

Para mais, descrevemos duas formas possíveis de encontrar um vetor básico admissível: num dos casos, supondo que temos à nossa disposição um vetor admissível arbitrário, e noutra supondo que o nosso pol provém de um pol na

forma canónica, com o vetor $b \geq 0$. Esta última condição, apesar de parecer restritiva, acaba por ser extremamente comum se pensarmos no pol no contexto económico.

Comecemos por finalizar a prova de completude do simplexo:

Prop 28. *O vetor \tilde{x} obtido após uma iteração do simplexo é sempre básico.*

Dem. Sejam a_1, \dots, a_n os vetores-coluna de A_{P_x} (que sabemos, por hipótese, formar uma base de \mathbb{R}^n) e seja \tilde{a} o vetor Ae_j .

Repare-se que os índices de coordenadas positivas de \tilde{x} são precisamente aqueles de x , mais o índice j , e retirando um ou mais dos índices correspondentes a a_1, \dots, a_n . Mais precisamente, se $\tilde{a} = \sum_{i=1}^n t_i a_i$, os índices retirados têm *necessariamente* o t_i correspondente diferente de zero, pois os índices associados a $t_i = 0$ não mudam de x para \tilde{x} .

Considere-se agora o que acontece quando, da lista a_1, \dots, a_n , se adiciona \tilde{a} e são retirados um ou mais vetores correspondentes a $t_i \neq 0$. A lista resultante é necessariamente linearmente independente: caso contrário, visto que os a_1, \dots, a_n são linearmente independentes, ter-se-ia que \tilde{a} se poderia escrever em função do resto dos vetores. Por estes formarem uma base, existe apenas uma forma de escrever \tilde{a} em função destes, e pelo menos um dos vetores necessários para isto foi removido da lista, pelo que entramos em contradição e a lista é linearmente independente. Isto prova que \tilde{x} é básico. \square

Isto permite-nos, então, afirmar:

Prop 29. *Se um pol é não-degenerado no sentido lato, a aplicação do algoritmo do simplexo termina sempre.*

É, então, útil assegurar que um pol dado é não-degenerado. A seguinte proposição dá um critério teórico para assegurar esta condição.

Prop 30. *Considere-se o seguinte pol:*

$$\begin{cases} \min_x cx \\ Ax = b \\ x \geq 0 \end{cases}$$

Onde a matriz A é $m \times n$. Dizemos que este é não-degenerado no sentido estrito se b não pode ser escrito como combinação linear de menos de m colunas de A .

Um pol não-degenerado no sentido estrito também o é no sentido lato.

Dem. Isto é, na realidade, uma trivialidade. Afinal, se houvesse um vetor básico degenerado x , b poderia ser escrito como combinação linear das (menos que m) colunas de A_{P_x} , pois $b = A_{P_x} x_{P_x}$. \square

Na prática, esta condição não é de particularidade utilidade. Tem, no entanto, duas aplicações teóricas:

- Em termos teóricos, é difícil categorizar o conjunto dos problemas não-degenerados no sentido lato. O conjunto dos problemas não-degenerados no sentido estrito, pelo outro lado, é muito mais fácil de manusear, visto que depende apenas de A . Por sua vez, isto permite-nos justificar em termos teóricos a afirmação ‘Problemas degenerados são raros’: de facto, é possível mostrar que, escolhida uma matriz A ‘ao calhas’, esta é degenerada (no sentido estrito) com probabilidade 0, o que implica que ela é não-degenerada no sentido lato com probabilidade 1. A prova disto não é aqui feita, visto que não está no âmbito da cadeira (e o autor não a sabe).
- A aplicação mais importante desta condição é, no entanto, os testes da Sernadas.

A Conceitos de Álgebra Linear

O leitor deverá estar familiar com os seguintes conceitos, definições e teoremas, que são aqui deixados sem prova. O leitor curioso poderá saciar a sua curiosidade com recurso a [1] e [2].

Definição. Sejam x, y pontos em \mathbb{R}^n . Define-se o produto interno de x e y , denotado $\langle x, y \rangle$, como $x_1y_1 + \dots + x_ny_n$. Se considerarmos os pontos em \mathbb{R}^n como sendo matrizes $n \times 1$, este produto pode ser escrito como $y^T x$, facto que será usado frequentemente.

Repare-se que, para qualquer x , $\langle x, x \rangle$ é sempre maior ou igual a zero, por ser uma soma de quadrados.

A norma de um ponto, $|x|$, é definida como $\sqrt{\langle x, x \rangle}$. A distância entre dois pontos, x, y , denotada $d(x, y)$, é definida como $|x - y|$.

Prop 31. (*Desigualdade de Cauchy-Schwarz*) $|\langle x, y \rangle| \leq |x||y|$
Igualdade acontece sse $x = \lambda y$ ou $y = \lambda x$, para algum $\lambda \in \mathbb{R}$.

Prop 32. (*Desigualdade triangular*) $|x + y| \leq |x| + |y|$

B Conceitos Topológicos

Definição. Seja $x \in \mathbb{R}^n$. Define-se a *vizinhança- ε de x* , denotada $V_\varepsilon(x)$, como o conjunto de pontos y tal que $d(x, y) < \varepsilon$. Implícito estará sempre que $\varepsilon > 0$.

Definição. Seja $S \subseteq \mathbb{R}^n$. Definem-se o interior, exterior, borda e fecho de S , respetivamente, da seguinte forma:

$$\text{int } S = \{ x \in S \mid \exists_\varepsilon V_\varepsilon(x) \subseteq S \}$$

$$\text{ext } S = \text{int}(\mathbb{R}^n \setminus S)$$

$$\partial S = \mathbb{R}^n \setminus \text{int } S \setminus \text{ext } S$$

$$\bar{S} = S \cup \partial S = \text{int } S \cup \partial S = \mathbb{R}^n \setminus \text{ext } S$$

Um conjunto S diz-se aberto se $S = \text{int } S$. Um conjunto S diz-se fechado se o seu complementar, $\mathbb{R}^n \setminus S$, for aberto, ou, equivalentemente, se $S = \bar{S}$.

Usando a desigualdade triangular é possível provar que $V_\varepsilon(x)$ é sempre aberto.

Definição. Seja $D \subseteq \mathbb{R}^n$, $f : D \rightarrow \mathbb{R}^m$. e seja $x \in D$. f diz-se *contínua em x* se:

$$\forall \varepsilon \exists \delta \forall y \in D \cap V_\delta(x) f(y) \in V_\varepsilon(f(x))$$

Uma definição alternativa e equivalente é a seguinte: Se x_n é uma sucessão de termos em D que converge para x , então $f(x_n)$ converge para $f(x)$.

Uma função $f : D \rightarrow \mathbb{R}^m$ diz-se *contínua* se for contínua em todos os pontos do seu domínio.

Prop 33. *Seja $D \subseteq \mathbb{R}^n$, $f : D \rightarrow \mathbb{R}^m$. Se f é contínua, a preimagem de qualquer aberto é aberta relativamente a D . Por outras palavras:*

Seja U um aberto contido em \mathbb{R}^m . Então, $f^{-1}(U)$ é da forma $D \cap V$, onde V é aberto em \mathbb{R}^m . Em particular, se f tiver domínio real, $f^{-1}(U)$ é aberto.

Mais ainda, isto é um sse. Ou seja, se esta última condição se verificar para qualquer aberto U , então f é contínua em todos os pontos.

Dem. Suponha-se, primeiro, que $f : D \rightarrow \mathbb{R}^m$ é contínua e que $U \subseteq \mathbb{R}^m$ é aberto. Mostre-se que $f^{-1}(U)$ é aberto relativamente a D .

Seja $x \in f^{-1}(U)$, ou seja, $f(x) \in U$. Visto que U é aberto, é deixado como exercício para o leitor pode observar que há uma bola aberta B_x à volta de x tal que para qualquer $x' \in D \cap B_x$ temos $f(x') \in U$. Sendo V a união destes B_x para $x \in D$, temos que V é aberto, por ser união de abertos; temos que para qualquer ponto x em $D \cap V$ se tem $f(x) \in U$, pelo que $D \cap V \subseteq f^{-1}(U)$, e finalmente, para qualquer ponto x em $f^{-1}(U)$ temos uma bola centrada em x contida em V , pelo que, em particular, x em $D \cap V$. Conclui-se, então, $D \cap V = f^{-1}(U)$, como desejado.

Prove-se agora a implicação oposta. Suponha-se que, para qualquer aberto U , $f^{-1}(U)$ é aberto relativamente a D . Mostre-se que f é contínua.

Fixe-se $x \in D$ e mostre-se que f é contínua em x . Escolha-se ε arbitrário. Seja B a bola de raio ε em torno de $f(x)$. Isto é aberto em \mathbb{R}^m , pelo que a sua preimagem, $f^{-1}(B)$ é, por hipótese, da forma $D \cap V$ com V aberto. $f(x)$ está contido em B , pelo que x está contido em $f^{-1}(B) \subseteq V$. Visto que este último é aberto, existe uma bola de raio δ em torno de x contida em V , pelo que, para $x' \in D \cap V_\delta(x)$ se tem $x' \in D \cap V$, donde, por definição de $D \cap V$, se tem $f(x') \in B = V_\varepsilon(f(x))$, como desejado. \square

A seguinte definição merece algumas palavras. O conceito prestes a ser definido é o de conjunto *compacto*. Isto é um conceito topológico, que se aplica a contextos altamente gerais. A definição geral é, no entanto, alto intragável.

Assim sendo, o leitor fica informado que a definição que se segue se aplica apenas a \mathbb{R}^n .

Definição. Seja $K \subseteq \mathbb{R}^n$. K diz-se compacto se é fechado e limitado.

Observação: o conceito de compacidade pode ser posto em termos de sucessões, em virtude do facto que tanto o conceito de conjunto fechado como o de conjunto limitado podem ser postos nestes termos. Temos a seguinte proposição:

Prop 34. *Um conjunto $S \subseteq \mathbb{R}^n$ é:*

- *Fechado sse qualquer sucessão convergente de elementos de S tem limite em S*
- *Limitado sse qualquer sucessão de elementos de S tem uma subsucessão convergente*
- *Compacto sse qualquer sucessão de elementos de S tem uma subsucessão com limite em S*

Dem. (Fechado implica sucessões convergentes convergem para dentro de S)
Por contradição: Suponha-se que S é fechado e que existe uma sucessão x_n de elementos de S cujo limite L não pertence a S . Assim sendo, visto que $L \in \mathbb{R}^n \setminus S$ e S é aberto, existe uma vizinhança ε de L que não intersesta S . Por definição de L , existe uma ordem a partir do qual todos os elementos de x_n estão em tal vizinhança. Em particular, existe pelo menos um elemento desta sucessão que está nesta vizinhança, e portanto não pertence a S , o que contradiz a hipótese sobre a sucessão.

(Sucessões convergentes convergem para dentro de S implica fechado) Por contra-recíproco: Suponha-se que S não é fechado. Existe então um ponto L na fronteira de S que não pertence a S . Visto que L está na fronteira de S , existem pontos de S arbitrariamente próximos de L . Escolha-se uma sucessão x_n de elementos de S tal que x_n esteja a menos de $\frac{1}{n}$ de L . Temos então por definição que x_n é uma sucessão de elementos de S que converge para L , limite este que não está em S .

(Limitado implica qualquer sucessão de elementos em S tem uma subsucessão convergente) Como o leitor pode facilmente verificar, para mostrar isto basta verificar que qualquer sucessão limitada tem uma subsucessão convergente. Para este efeito, seja x_n uma sucessão limitada.

Por ser limitada, esta sucessão está contida num retângulo fechado R . Considere-se o seguinte processo: visto que a sucessão é infinita, existem infinitos n tal que $x_n \in R$. (Em particular, todos). Divida-se R em 2^n retângulos idênticos fechados (partindo-o ao meio ao longo de cada eixo). Visto que há infinitos termos, pelo menos um destes retângulos contém um número infinito de termos. Chame-se a tal retângulo R_1 . Repita-se o processo com R_1 para obter um retângulo R_2 que contém um número infinito de termos, e assim por diante. Visto que a sequência de retângulos obtida é uma sequência descendente de

retângulos fechados, a interseção contém pelo menos um ponto L . É agora fácil verificar que, se se definir i_1 como sendo um índice tal que $x_{i_1} \in R_1$, i_2 como sendo um índice maior que i_1 tal que $x_{i_2} \in R_2$, e assim por diante, temos que a subsequência de x_n dada por x_{i_n} converge para L . Esta verificação é deixada ao leitor.

(Qualquer sucessão de elementos de S tem subsequência convergente implica S limitado) Por contra-recíproco: Suponha-se que S é ilimitado, e seja então x_n uma sucessão de elementos de S tal que $|x_n| > n$. É fácil verificar que qualquer subsequência de x_n satisfaz esta propriedade, pelo que não converge. Logo, x_n não tem subsequências convergentes.

Provados estes dois resultados, o resultado para compacidade é trivial. \square

Como consequência, tem-se o seguinte teorema, que o leitor poderá interpretar como uma generalização do teorema de Weierstrass.

Prop 35. *Funções contínuas levam compactos em compactos.*

Dem. Seja $f : D \rightarrow \mathbb{R}^m$ contínua, com $D \subseteq \mathbb{R}^n$ um compacto. Seja $S = f(D)$. Para mostrar que S é compacto, fixe-se uma sucessão y_n de elementos de S e encontre-se uma subsequência convergente desta.

Sabendo que $y_n \in f(D)$ para todo n , existe uma sucessão $x_n \in D$ tal que $f(x_n) = y_n$. Ora, visto que a sucessão x_n está contida no compacto D , existe uma subsequência x_{i_n} que converge para algum $x \in D$. Assim sendo, é fácil ver que a sucessão $y_{i_n} = f(x_{i_n})$ converge para $f(x) \in f(D)$.

Temos a demonstração concluída, visto que começámos com uma sucessão arbitrária de elementos de S e encontramos uma subsequência convergente desta cujo limite está em S . \square

Prop 36. *Seja $T : \mathbb{R}^n \rightarrow \mathbb{R}^m$ uma transformação linear. Então, T é contínua.*

Dem. Seja x_n uma sucessão de elementos de \mathbb{R}^n de limite L . Mostre-se que a sucessão $T(x_n)$ tem limite $T(L)$. Isto é equivalente a provar que a sucessão $T(x_n - L)$ tem limite 0. Para isso, é suficiente mostrar que se $x_n \rightarrow 0$ então $T(x_n)$ converge para $T(0)$.

Considere-se agora o seguinte: suponha-se momentaneamente que x_n não tem nenhum termo igual a zero. Então, $T(x_n) = |x_n|T(\frac{x_n}{|x_n|})$, ou seja, $|x_n|$ vezes $T(z_n)$, onde z_n é um elemento do círculo unitário. (É possível levantar agora a restrição $x_n \neq 0$, visto que esta última afirmação continua a ser verdade para $x_n = 0$; basta escolher z_n arbitrário.)

Repare-se agora no seguinte: o círculo unitário é compacto. Logo, a imagem do círculo sob T é compacta, pelo que, em particular, é limitada. Logo, existe C tal que, para z unitário, $|T(z)| \leq C$, donde $|T(x_n)| = |x_n|T(z_n) \leq C|x_n|$. Ora, como $x_n \rightarrow 0$ temos que este último termo vai para zero, donde $T(x_n)$ vai para zero, como desejado. \square

Referências

- [1] Linear Algebra Done Right, Sheldon Axler
- [2] Calculus on Manifolds, Michael Spivak
- [3] Foundations and Applications of Linear Optimization, Cristina Sernadas et al.