Programação Linear

Marco Antonio Figueiredo Menezes

Departamento de Computação da
Universidade Católica de Goiás (UCG)
Endereço Eletrônico: marco@ucg.br

Setembro/2006

[0] Este texto está em elaboração. A primeira versão deste texto foi revisada pelo Professor Clóvis Gonzaga. Uma segunda versão foi revisada pelo Professor Hilton Machado. Esta é a terceira versão. É importante lembrar que desde o início, a partir de 1999, o texto foi utilizado e corrigido em minicursos (IV e V Encontro de Matemática e Estatística da UFG, III UCGComp e V UCGComp e IV Encontro de Matemática da UEG-Unidade Anápolis), pelos meus alunos de Iniciação Científica (1999-2004), no Seminários de Otimização que coordenamos aqui no Departamento de Computação da UCG desde 1999 e na disciplina Análise Aplicada Computacional (CMP4132) nas turmas de Ciência da Computação e Engenharia de Computação da UCG (1999/2-2005/1).

Sumário

Capítulo 1

Um Pouco da História

Introduziremos a Programação Linear (PL) nos três primeiros capítulos. O que faremos neste caminho, então, será reescrever o que já existe na literatura, comentando alguns resultados que julgamos serem importantes para a construção da história da PL, definindo o problema de PL e modelando problema(s) de Programação Linear (PPL). Iniciamos descrevendo um pouco da história da Programação Linear.

A PL poderia ter sido iniciada em torno de 1758 quando os economistas começaram a descrever sistemas econômicos em termos matemáticos. Também, Walras propôs em 1874 um sofisticado modelo matemático que tinha como parte da sua estrutura coeficientes tecnológicos fixados.

O famoso matemático Fourier parece ter sido o primeiro a estudar desigualdades lineares para a Mecânica e para a Teoria das Probabilidades. Ele estava interessado em encontrar o ponto mínimo em um poliedro. Ele sugeriu uma solução por uma descida de vértice em vértice para um mínimo, que é o princípio por trás do método simplex desenvolvido por Dantzig. Este é provavelmente o primeiro exemplo, datado de 1826, de um problema de PL. Mais tarde, em 1911, outro matemático famoso, Poussin, considerou o mesmo problema e propôs uma solução similar. Veja página 21 em Dantzig [14].

Em 1939, o matemático e economista Kantorovich, formulou e resolveu um problema de PL tratando com organização e planejamento de produção. Devido a guerra fria entre os Estados Unidos da América e a então União Soviética, este trabalho ficou desconhecido para o Ocidente durante uns vinte anos.

A literatura matemática continha inúmeros artigos concernentes a técnicas para resolver sistemas de equações lineares. Por outro lado, o estudo de sistemas de desigualdades lineares não despertava interesse até o advento da Teoria dos Jogos em 1944 e de PL em 1947.

O problema geral de PL foi primeiramente desenvolvido e aplicado em 1947 por Dantzig, Wood e seus associados no Departamento da Força Aérea dos Estados Unidos. Este grupo foi solicitado para pesquisar a viabilidade em aplicar a Matemática e técnicas relacionadas para os problemas de planejamento e programação militar. Mais tarde, em outubro de 1948, este grupo recebeu o título oficial de Projeto SCOOP (Scientific Computation of Optimum Programs). O artigo fundamental circulou confidencialmente por alguns anos e foi publicado por Dantzig [13] em 1951.

A influência da Segunda Guerra Mundial foi decisiva para o surgimento da PL e seu posterior desenvolvimento. Haviam necessidades (planejar, transportar) e financiamentos (o Projeto SCOOP; o desenvolvimento dos computadores; a Conferência de Chicago em 1949, onde matemáticos, economistas e estatísticos de instituições acadêmicas e de várias agências governamentais apresentaram trabalhos usando PL).

Em Teoria da Computação, um algoritmo é considerado eficiente quando o seu número de passos for limitado polinomialmente, conforme Edmonds [17] e Cobham [12] (a distinção entre algoritmos polinomiais e exponenciais já havia sido sugerida por von Neumann [60]). Em 1972, Klee e Minty [31] construíram um exemplo para estabelecer a não polinomialidade do algoritmo simplex para um certo critério de escolha para a entrada na base. A partir daí, vários outros exemplos inviabilizaram novos critérios de escolha (veja Shamir [53]). Algoritmos do tipo elipsóide foram introduzidos por Shor [54] e Yudin e Nemirovskii [63] para Programação Convexa. Em 1979 e 1980, Khachiyan [29] e [30], respectivamente, utilizou o método dos elipsóides para o problema de viabilidade de PL com dados inteiros. Khachiyan definiu o número L e mostrou que seu algoritmo resolve o PPL em tempo polinomial. Todavia, no final da década de 70, convivemos com um fato curioso em PL: por um lado, o método simplex com complexidade exponencial, mas que funciona bem na prática e, por outro lado, o método dos elipsóides com complexidade polinomial, mas que funciona mal na prática.

Em 1984, Karmarkar [28] publicou seu algoritmo de pontos interiores baixando a complexidade em relação ao método de Khachiyan. Ele obteve um limite para o número de iterações de $O(nL)$ e um número de operações aritméticas por iteração de $O(n^{2,5})$, totalizando $O(n^{3,5}L)$ operações aritméticas. A trajetória central foi inicialmente estudada por Bayer e Lagarias [4] e

Megiddo [40]. Em 1986, Renegar [51] provou que o método de centros de Huard escrito em termos da função barreira logarítmica é polinomial para problemas de PL, se o problema de minimização auxiliar é tratado pelo método de Newton (Veja Nesterov [47]). Renegar obteve um limite de complexidade de $O(\sqrt{n}L)$ iterações, mas com complexidade total igual à de Karmarkar. Em 1987, Gonzaga [22] e Vaidya [57] obtiveram simultaneamente algoritmos com a complexidade de $O(n^3L)$, onde o primeiro desenvolve um algoritmo que usa uma função de penalidade do tipo barreira logarítmica e, o segundo, seguindo a mesma metodologia de Renegar. Em 1999, Anstreicher [1] obteve a complexidade de $O((n^3/\ln(n))L)$.

Os algoritmos de ponto-interior-inviável primais-duais para PL são considerados os algoritmos de pontos interiores mais eficientes na prática (veja Lustig, Marsten e Shanno [34] e, veja também, Marsten, Subramanian, Saltzman, Lustig e Shanno [38]). Kojima, Megiddo e Mizuno [32] provaram a convergência de um algoritmo primal-dual de ponto-interior-inviável para PL. Zhang [64] demonstrou convergência polinomial em $O(n^2L)$ iterações detectando inviabilidade para uma determinada região. Mizuno [43] demonstrou $O(n^2L)$ para o algoritmo de Kojima, Megiddo e Mizuno e desenvolveu uma variante deste obtendo $O(nL)$ iterações. Ambos algoritmos detectam inviabilidade para uma determinada região. A variante de Mizuno foi um algoritmo preditor-corretor. Quase simultaneamente, Potra [48] obteve a mesma complexidade em iterações. Ainda, Potra [49] demonstrou convergência quadrática e $O(\sqrt{n}L)$ iterações sob determinadas condições para o ponto inicial. Ye, Todd e Mizuno [62] desenvolveram um método homogêneo e auto-dual para PL, que possui $O(\sqrt{n}L)$ iterações para uma formulação artificial do problema de PL. Todavia, no momento, a menor complexidade em iterações para a formulação não artificial é $O(nL)$. Supondo viabilidade dual, Ye [61] obtém $O(\sqrt{n}L)$ iterações para métodos de redução potencial combinando fases 1 e 2 e Menezes e Gonzaga [42] obtêm a mesma complexidade relacionando parâmetros de viabilidade e otimalidade com passos curtos.

Uma experiência para resolver um PPL com 837 restrições e 12.753.313 variáveis pode ser encontrada no artigo de Bixby, Gregory, Lustig, Marsten e Shanno [8], os quais tratam com um algoritmo híbrido ponto interior/simplex, específico para um problema que aparece no planejamento de tripulação aérea. De fato, este trabalho fornece uma alternativa relevante no tratamento de PPL de grande porte.

No próximo capítulo definiremos o problema de Programação Linear, o qual é o problema que nos propomos estudar.

Capítulo 2

O Problema

Continuamos com a nossa introdução à Programação Linear. Nosso objetivo aqui é definir o problema de PL. Em particular, definimos o PPL na forma padrão, no sentido de que qualquer PPL pode ser convertido para este formato.

2.1 O problema no formato padrão

A Otimização pode ser vista como uma área da Matemática Aplicada, em Matemática; ou uma área da Pesquisa Operacional, em Engenharias; ou uma área da Matemática Computacional, em Computação. Independentemente da área, notamos que a Otimização é uma sub-área interdisciplinar e importante.

O problema de Otimização consiste em encontrar, se possível, os minimizadores (ou os maximizadores) de uma função definida em uma determinada região.

Consideremos os números inteiros m e n tais que $n > m > 0$. Dados uma matriz numérica com coeficientes reais A, $m \times n$, e vetores $b \in R^m$ e $c \in R^n$, o problema de Programação Linear no formato padrão é o seguinte problema de Otimização, usualmente denominado problema primal:

$$\begin{array}{lll} (P) & \textit{minimizar} & c^T x \\ & \textit{sujeito a:} & Ax = b \\ & & x \geq 0. \end{array}$$

Seguem-se algumas definições associadas ao problema (P).

Definição 2.1.1 *Considere o PPL (P).*

(a) A função linear $x \mapsto c^T x$ é chamada função objetivo.

(b) O conjunto

$$\mathcal{X} = \{x \in R^n; Ax = b, x \geq 0\}$$

é chamado conjunto viável e um ponto $x \in \mathcal{X}$ é denominado ponto viável.

(c) O conjunto $\{x \in R^n; x > 0\}$ é chamado conjunto de pontos interiores e um ponto deste conjunto é denominado ponto interior. O conjunto $\mathcal{X}^0 = \{x \in \mathcal{X}; x > 0\}$ é chamado conjunto de pontos interiores viáveis e um ponto $x \in \mathcal{X}^0$ é denominado ponto interior viável.

(d) Quando existe, o número $v(P) = \min\{c^T x; x \in \mathcal{X}\}$ é denominado o valor ótimo ou custo ótimo. O conjunto

$$\mathcal{X}(P) = \{x \in \mathcal{X}; c^T x = v(P)\}$$

é chamado conjunto de soluções ótimas e um ponto $x \in \mathcal{X}(P)$ é denominado solução ótima ou minimizador ou ponto de mínimo.

(e) O problema (P) chama-se problema ilimitado, quando existe uma seqüência (x^k) tal que $x^k \in \mathcal{X}$ e $c^T x^k \to -\infty$, quando $k \to \infty$.

(f) O problema (P) chama-se problema inviável, quando $\mathcal{X}$ é vazio.

Podemos interpretar o PPL (P) da seguinte maneira: dados uma matriz tecnológica com números reais A, $m \times n$, um vetor do lado direito $b \in R^m$ e um vetor custo $c \in R^n$, encontrar, se existir, um (e não 'o') vetor de variáveis de decisão $x^* \in \mathcal{X}$ tal que

$$c^T x^* = \min\{c^T x; x \in \mathcal{X}\}.$$

Atenção: Que tal fazer o exercício 1 agora?

Geralmente, pretendemos resolver um PPL no formato do problema (P). Isto é, o primeiro grupo de restrições envolve somente igualdades e todas as variáveis do modelo são não negativas. Além disso, queremos minimizar o valor da função objetivo.

Neste sentido, como reduzir um PPL qualquer para o formato padrão? A resposta vem a seguir, iniciando com uma proposição, cuja demonstração pode ser encontrada nas páginas 67 e 68 em [11].

Proposição 2.1.2 *Seja C um conjunto viável de um PPL que possui alguma solução ótima. Então,*

$$\max\{c^T x; x \in C\} = -\min\{-c^T x; x \in C\}.$$

Demonstração: Seja $x^* \in C$ uma solução ótima do problema

$$\max\{c^T x; x \in C\}.$$

Para todo $x \in C$ temos: pela definição de x^*, $\max\{c^T x\} = c^T x^* \geq c^T x$, que é equivalente a, $-c^T x \geq -c^T x^* = \min\{-c^T x\}$. Portanto, para todo $x \in C$, $\max\{c^T x\} = -\min\{-c^T x\}$, finalizando a demonstração. ∎

Na ocorrência de

$$\begin{array}{ll} \textit{maximizar} & c^T x \\ \textit{sujeito a:} & x \in \mathcal{X}, \end{array}$$

podemos usar a proposição anterior assim,

$$\begin{array}{ll} \textit{minimizar} & -c^T x \\ \textit{sujeito a:} & x \in \mathcal{X}. \end{array}$$

Observe que, caso exista solução ótima, dito $x^* \in \mathcal{X}$, o valor ótimo será $\max\{c^T x; x \in \mathcal{X}\} = -\min\{-c^T x; x \in \mathcal{X}\}$, pois $c^T x^* = -(-c^T x^*)$.

Atenção: Que tal fazer o exercício 2(a) agora?

Considere $i = 1, \ldots, m$. Na ocorrência de desigualdades do tipo

$$\sum_{j=1}^{n} a_{ij} x_j \leq b_i \text{ ou } \sum_{j=1}^{n} a_{ij} x_j \geq b_i,$$

basta tomarmos $x_{n+i} \geq 0$ tal que

$$\sum_{j=1}^{n} a_{ij}x_j + x_{n+i} = b_i \text{ ou } \sum_{j=1}^{n} a_{ij}x_j - x_{n+i} = b_i,$$

respectivamente. Dizemos que x_{n+i} é uma variável de folga, quando adicionada na restrição e, variável de excesso, quando subtraída na restrição.

Atenção: Que tal fazer o exercício 2(b) agora?

Por outro lado, considere $j = 1, \ldots, n$. Na ocorrência de variáveis livres, quer dizer, sem restrição de sinal, isto é,

$$x_j \in R,$$

basta realizarmos uma mudança de variáveis definindo,

$$x_j = \bar{x}_j - \hat{x}_j \text{ com } \bar{x}_j \geq 0 \text{ e } \hat{x}_j \geq 0.$$

Atenção: Que tal fazer o exercício 2(c) agora?

Ainda, considere $j = 1, \ldots, n$. Sejam dados números reais l_j e u_j, com $l_j \neq 0$. Na ocorrência de variáveis do tipo

$$x_j \geq l_j \text{ ou } x_j \leq u_j,$$

podemos considerar $x_j \geq l_j$ ou $x_j \leq u_j$ como restrições do tipo $\geq$ ou $\leq$, respectivamente. Assim, aumentamos o número de restrições m para $m + 2n$ e o número de variáveis de n para $3n$. Como uma segunda alternativa, podemos definir x_j por $\bar{x}_j + l_j$, para $\bar{x}_j \geq 0$ e $x_j \leq u_j$ como restrições do tipo $\leq$. Assim, aumentamos o número de restrições m para $m + n$ e o número de variáveis de n para $2n$. Ainda, para $x_j \geq l_j$ e $x_i \leq u_i$ com $j \neq i$, $i = 1, \ldots, n$, podemos definir

$$x_j = \bar{x}_j + l_j, \ \bar{x}_j \geq 0 \text{ e } x_i = u_i - \bar{x}_i, \ \bar{x}_i \geq 0.$$

Assim, mantemos o número de restrições e variáveis e alteramos o vetor do lado direito. Finalmente, se tivermos

$$l_j \leq x_j \leq u_j, \text{ isto é, } x_j \geq l_j \text{ e } x_j \leq u_j,$$

existem técnicas eficientes para a resolução de problemas de PL com estas restrições denominadas restrições tipo caixa (veja Bazaraa, Jarvis e Sherali

[5]). Todavia, neste último caso (para n "pequeno"), podemos usar a primeira ou a segunda alternativas.

Note que não consideramos desigualdades estritas.

Atenção: Que tal fazer os exercícios 2 e 3 agora?

Sem perda de generalidade, supomos $posto(A) = m$; isto é, a matriz A possui m linhas (ou colunas) linearmente independentes, e dizemos que A possui posto completo. Com efeito, se as linhas da matriz A são linearmente dependentes, então (P) é inviável ou as equações lineares redundantes podem ser removidas seqüencialmente até a matriz resultante ter posto completo (posto máximo).

Além disso, se A é uma matriz nula e b é um vetor nulo, o posto de A é igual a zero que é menor do que m. Neste caso, minimizar o valor de uma função objetivo sujeito ao primeiro ortante (quadrante no R^2, octante no R^3, etc.), possui o vetor nulo como uma solução, quando o vetor $c \geq 0$, ou caso contrário, trata-se de um problema ilimitado. Se $b \neq 0$ (para A nula), então trata-se de um problema inviável.

Atenção: Que tal fazer os exercícios 4 e 5 agora?

A propósito, quando um problema de Otimização é um problema de PL? Quando as funções envolvidas (a função objetivo e as restrições do problema) são afins (lineares) e contínuas.

No próximo capítulo introduziremos o processo de modelagem em PL, que evidenciará a relevância prática para o estudo dos métodos e dos algoritmos para resolver PPL.

Atenção: Que tal fazer os exercícios 6 e 7 agora?

2.2 Exercícios

1. Para o problema

$$\begin{array}{lll} min & -x_1 & \\ s.\ a: & x_1 + x_2 & = 2 \\ & x_1,\quad x_2 & \geq 0, \end{array}$$

identifique:

(a) a matriz tecnológica A, 1×2, o vetor do lado direito $b \in R$ e o vetor custo $c \in R^2$; e

(b) a função objetivo; os conjuntos: viável, de pontos interiores, de pontos interiores viáveis e de soluções ótimas; alguns pontos: viáveis, interiores, interiores viáveis e a solução ótima; e o valor ótimo.

2. Colocar os seguintes problemas de PL no formato padrão:

(a)
$$\begin{array}{lll} \textit{maximizar} & x_1 & \\ \textit{sujeito a}: & x_1 + x_2 & = 2 \\ & x_1, \quad x_2 & \geq 0. \end{array}$$

(b)
$$\begin{array}{lll} \textit{minimizar} & x_1 & \\ \textit{sujeito a}: & x_1 + x_2 & \leq 2 \\ & x_1 + x_2 & \geq 1 \\ & x_1, \quad x_2 & \geq 0. \end{array}$$

(c)
$$\begin{array}{lll} \textit{minimizar} & x_1 & \\ \textit{sujeito a}: & x_1 + x_2 & = 2. \end{array}$$

(d)
$$\begin{array}{lll} \textit{minimizar} & x_1 & \\ \textit{sujeito a}: & x_1 + x_2 & = 2 \\ & x_1 \geq -1, \quad x_2 \leq 0. & \end{array}$$

(e)
$$\begin{array}{lll} \textit{maximizar} & \sqrt{2}x_1 - x_2 & \\ \textit{sujeito a}: & x_1 + 0,005x_2 + x_3 & \leq 3000 \\ & x_2 - x_3 & \geq 1 \\ & x_2 \geq 0, \quad 0,1 \leq x_3 \leq 8. & \end{array}$$

3. Considere o PPL do item (c) do exercício anterior. Quantas variáveis você obteve para o formato padrão? Você consegue ficar com apenas três? Se sim, como? Senão, tente! Observe que para n variáveis livres, sempre podemos realizar uma mudança de variáveis de tal maneira que obtemos $n + 1$ variáveis não negativas, ao invés de $2n$, por exemplo. Dependendo da situação, isto pode ser vantajoso computacionalmente.

4. Sejam dados a matriz A, 2×3, e os vetores $b \in R^2$ e $c \in R^3$, a saber:

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}, \; b = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \text{ e } c = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

   Pede-se:

   (a) Qual é o posto da matriz A? A tem posto completo?

   (b) Considere o PPL no formato padrão definido por A, b e c. Este problema é um problema inviável ou as equações lineares redundantes podem ser removidas seqüencialmente até a matriz resultante ter posto completo? Justifique.

   (c) Considere a questão do item anterior. E se $b = [2, 2]^T$?

5. Considere o PPL (P) com $posto(A) = m$. Verifique que a hipótese $n > m > 0$ não é restritiva para o problema (P). Quer dizer, verifique que: se $m = n$, o sistema de equações lineares possui uma única solução $\hat{x}$. Se $\hat{x} \geq 0$, então $\hat{x}$ é a única solução ótima para o problema (P). Senão, (P) é um problema inviável.

6. Quais dos seguintes problemas de Otimização são problemas de PL? Justifique.

   (a)

$$\begin{array}{ll} minimizar & \sqrt{5}x_3 \\ sujeito\ a: & x_1 - x_2 + x_3 \ \leq 1 \\ & x_2 \leq 0. \end{array}$$

   (b)

$$\begin{array}{lll} minimizar & \log(x_1 + 1) & \\ sujeito\ a: & x_1 + 2x_2^2 & = 2 \\ & x_1, \ \ x_2 & \geq 0. \end{array}$$

   (c)

$$\begin{array}{ll} minimizar & -x_1 \\ sujeito\ a: & x_1 - x_2 = 0 \\ & x_1, \ \ x_2 \in \{0, 1\}. \end{array}$$

7. **(Página 12 em [18])** Considere o seguinte problema de Otimização:

$$\begin{array}{ll} \textit{minimizar} & |x_1| \\ \textit{sujeito a}: & x_1 + x_2 \quad = 7 \\ & x_2 \geq 0. \end{array}$$

Converta este problema de Programação Não Linear, não diferenciável, para um PPL no formato padrão, usando a seguinte sugestão: para qualquer número real x, pode-se encontrar $u, v \geq 0$, tais que $|x| = u + v$ e $x = u - v$. A propósito, os valores de u e v são únicos.

Capítulo 3

Sobre a Modelagem

Neste capítulo, finalizaremos a nossa introdução à PL. Nosso objetivo aqui é introduzir o processo de modelagem para PPL, em contraposição ao uso apenas da experiência e do bom senso. Neste capítulo não estamos interessados em resolver problemas de PL.

Citamos os primeiro e segundo capítulos do livro de Goldbarg e Luna [19] como uma boa referência para o estudo de modelagem.

Os responsáveis pela tomada de decisões nos mais variados campos da atividade humana defrontam-se com a necessidade de resolver algum problema específico de Otimização. O primeiro passo fundamental é a formulação, que consiste em traduzir a realidade empírica para o estabelecimento do modelo matemático. No entanto, a correspondência entre sistema real e modelo formal está longe de ser perfeita: a tradução está sujeita a erros, simplificações e falhas de comunicação. Também, não existem técnicas precisas capazes de permitir o estabelecimento do modelo de um problema. Para consegui-lo, é importante a nossa capacidade de análise e síntese. O próximo passo é a dedução do modelo, isto é, analisá-lo e resolvê-lo através de algoritmos específicos. Sua solução, atenta aos métodos numéricos em Computação, sugere uma tomada de decisão. Para a sua sustentação, recorremos ao próximo passo que é a interpretação de uma solução do modelo para uma solução do sistema real. Neste ponto, o uso da experiência e do bom senso é de significativa relevância. Se o modelo não for validado, ele deve ser reformulado, e assim por diante. Este é o processo de modelagem (veja Ravindran, Phillips e Solberg [50]).

Considerando o processo de modelagem, estudaremos neste capítulo o passo da formulação.

A seguir vamos enunciar um problema da dieta. O problema da dieta é famoso na literatura da PL porque ele foi o primeiro problema econômico resolvido, em princípio como teste, após o advento da disciplina PL (veja Namen e Bornstein [45]).

3.1 Um problema da dieta

Nesta seção, estudaremos um problema da dieta em nutrição de ruminantes para Programação Linear. Para uma abordagem em Programação Não Linear, veja Menezes e Vieira [41].

Vamos separar esta seção em duas subseções: o problema prático e o modelo matemático.

3.1.1 O problema

Suponhamos que fomos convidados para resolver o problema de minimizar o custo de uma dieta para vacas leiteiras em lactação para uma produção de leite de 20kg/dia por vaca.

Um cientista em nutrição de ruminantes[1] sugere que, para uma boa dieta, a vaca em lactação deve consumir feno, silagem de milho, fubá, farelo de soja e mistura mineral.

Para cada quilo de feno, temos em média 50g de proteína, 1,2Mcal de energia, 750g de fibra, 10g de cálcio e 1g de fósforo. Para cada quilo de silagem de milho, temos em média 30g de proteína, 1,5Mcal de energia, 600g de fibra, 6g de cálcio e 1g de fósforo. Para cada quilo de fubá, temos em média 90g de proteína, 2,8Mcal de energia, 90g de fibra, 0,9g de cálcio e 0,2g de fósforo. Para cada quilo de farelo de soja, temos em média 450g de proteína, 2,5Mcal de energia, 150g de fibra, 3,2g de cálcio e 1,6g de fósforo. Para cada quilo de mistura mineral, temos em média 0g de proteína, 0Mcal de energia, 0g de fibra, 320g de cálcio e 160g de fósforo.

Além disso, para uma boa dieta, a vaca em lactação deve possuir na sua alimentação pelo menos 2240g/dia de proteína, pelo menos 24Mcal/dia de energia, pelo menos 6000g/dia de fibra e no mínimo 64g/dia e no máximo 160g/dia de cálcio e fósforo.

[1] Os dados que se seguem foram fornecidos pelo colega zootecnista Ricardo Vieira, quando no Instituto Melon, através de uma comunicação telefônica.

O preço médio do quilo de feno, silagem de milho, fubá, farelo de soja e mistura mineral são, respectivamente, 35 centavos, 30 centavos, 20 centavos, 28 centavos e 60 centavos.

3.1.2 O modelo

Uma vez que já dissemos não haver regras precisas para o processo de modelagem, sugerimos uma tentativa de encontrar inicialmente as variáveis de decisão. E, finalmente, sugerimos verificar as unidades de grandeza de cada dado, logo, das variáveis de decisão também.

Neste caso, definimos x_j, $j = 1, 2, \ldots, 5$, as variáveis de decisão que pretendemos encontrar, se existir, a saber:

x_j : quantidade em quilogramas por dia do j-ésimo alimento para cada vaca.

Quer dizer, em quilogramas, para $j = 1$ temos a quantidade do alimento feno, para $j = 2$ temos a quantidade do alimento silagem de milho, para $j = 3$ temos a quantidade do alimento fubá, para $j = 4$ temos a quantidade do alimento farelo de soja e para $j = 5$ temos a quantidade do alimento mistura mineral.

Aqui, o nosso objetivo é minimizar o custo da compra dos alimentos, a saber:

$$0,35x_1 + 0,30x_2 + 0,20x_3 + 0,28x_4 + 0,60x_5.$$

Nosso objetivo de minimização está sujeito a algumas restrições. Sabemos que cada quilo de feno fornece 10g de cálcio, cada quilo de silagem de milho fornece 6g de cálcio, cada quilo de fubá fornece 0,9g de cálcio, cada quilo de farelo de soja fornece 3,2g de cálcio e cada quilo de mistura mineral fornece 320g de cálcio. Assim, temos que x_1 quilos de feno, x_2 quilos de silagem de milho, x_3 quilos de fubá, x_4 quilos de farelo de soja e x_5 quilos de mistura mineral fornecerão conjuntamente $10x_1 + 6x_2 + 0,9x_3 + 3,2x_4 + 320x_5$g/dia de cálcio, um total que deve atender o mínimo desejado de 64g/dia de cálcio e o máximo desejado de 160g/dia de cálcio. Desta forma, podemos construir duas inequações, a saber:

$$10x_1 + 6x_2 + 0,9x_3 + 3,2x_4 + 320x_5 \geq 64$$

e

$$10x_1 + 6x_2 + 0,9x_3 + 3,2x_4 + 320x_5 \leq 160.$$

Da mesma maneira, obtemos as seguintes desigualdades:

$$\begin{gathered} 50x_1 + 30x_2 + 90x_3 + 450x_4 \geq 2240 \\ 1,2x_1 + 1,5x_2 + 2,8x_3 + 2,5x_4 \geq 24 \\ 750x_1 + 600x_2 + 90x_3 + 150x_4 \geq 6000 \\ x_1 + x_2 + 0,2x_3 + 1,6x_4 + 160x_5 \geq 64 \\ x_1 + x_2 + 0,2x_3 + 1,6x_4 + 160x_5 \leq 160. \end{gathered}$$

Finalmente, compramos os alimentos ou não. Quer dizer, não podemos ter quantidades negativas de alimentos. Então,

$$x_j \geq 0, \ j = 1, 2, \ldots, 5.$$

Portanto, o nosso modelo matemático que tenta traduzir uma particular realidade na dieta de vacas leiteiras em lactação com uma produção de leite de 20kg/dia por vaca com um gasto mínimo na compra dos alimentos, é dado pelo PPL

$$\begin{array}{lcl} \textit{minimizar} & 0,35x_1 + 0,30x_2 + 0,20x_3 + 0,28x_4 + 0,60x_5 & \\ \textit{sujeito a:} & 50x_1 + 30x_2 + 90x_3 + 450x_4 & \geq 2240 \\ & 1,2x_1 + 1,5x_2 + 2,8x_3 + 2,5x_4 & \geq 24 \\ & 750x_1 + 600x_2 + 90x_3 + 150x_4 & \geq 6000 \\ & 10x_1 + 6x_2 + 0,9x_3 + 3,2x_4 + 320x_5 & \geq 64 \\ & 10x_1 + 6x_2 + 0,9x_3 + 3,2x_4 + 320x_5 & \leq 160 \\ & x_1 + x_2 + 0,2x_3 + 1,6x_4 + 160x_5 & \geq 64 \\ & x_1 + x_2 + 0,2x_3 + 1,6x_4 + 160x_5 & \leq 160 \\ & x_1, \quad x_2, \quad x_3, \quad x_4, \quad x_5 & \geq 0. \end{array}$$

No próximo capítulo caracterizaremos o conjunto viável como um conjunto poliedral, iniciando uma segunda etapa: os fundamentos da PL. Assim, neste momento, estamos finalizando a etapa introdutória.

3.2 Exercícios

1. **(Página 399 em [10])** A companhia Sovina de Investimentos possui seis milhões de reais, quantia esta que deverá ser aplicada em 5 tipos de

investimentos, sendo que os retornos anuais para cada investimento são: investimento 1 (I1), 10%; investimento 2 (I2), 8%; investimento 3 (I3), 6%; investimento 4 (I4), 5%; investimento 5 (I5), 9%. O gerente desta companhia deseja diversificar os investimentos para obter o máximo rendimento possível. Dado o elemento de risco envolvido, o gerente restringiu a quantia aplicada em I1 a não mais que a quantia total aplicada em I3, I4 e I5 (em conjunto). A soma da quantia total aplicada em I2 e I5 deve ser pelo menos igual à quantia aplicada em I3. O I2 deve estar limitado a um nível que não exceda a quantia aplicada em I4. É preciso determinar a alocação ótima de investimento entre as cinco categorias, de forma que o retorno ao final do ano seja o máximo possível. Formular somente o problema.

2. **(Página 400 em [10])** A companhia ZigZag possui fábricas em Campinas e Belo Horizonte (BH). Esta companhia produz e distribui computadores a comerciantes de várias cidades. Numa determinada semana, a companhia possui: 30 unidades em Campinas e 40 unidades em BH. Nesta mesma semana, esta companhia deve atender os pedidos dos comerciantes das seguintes cidades: 20 unidades para São Paulo (SP), 25 unidades para o Rio de Janeiro (RJ) e 25 unidades para Vitória. O problema consiste em distribuir as máquinas aos comerciantes de forma a atender os pedidos a um custo mínimo de transporte. Os custos unitários de transporte em reais são: 9 de Campinas para SP, 16 de Campinas para RJ, 25 de Campinas para Vitória, 27 de BH para SP, 22 de BH para RJ e 21 de BH para Vitória. Qualquer quantidade fracionária é aceitável. Formular somente o problema.

3. **(Página 77 em [11])** Numa determinada região, o problema da poluição atmosférica tornou-se crítico devido à emissão do poluente SO_2 por um certo número n de fábricas. Este poluente é liberado pela queima de m combustíveis para a produção da energia necessária. Cada fábrica j necessita diariamente e_j unidades de energia. Cada tonelada do combustível i, cujo custo é c_i, produz a_{ij} unidades de energia e emite p_{ij} unidades do poluente, quando utilizada na fábrica j. Deseja-se que a emissão diária do poluente para a região não exceda α unidades. Por uma questão de eqüidade na distribuição dos custos da poluição, é importante assegurar adicionalmente que o custo da unidade de energia produzida seja o mesmo para as n fábricas. Pretende-se minimizar o

custo total de energia para as n fábricas. Formular o problema como um problema de PL.

4. **(Página 12 em [18])** Uma empresa chamada CHIPCO produz dois tipos de chips de memória para computadores. O preço unitário de venda são 15 reais para o chip 1 e 25 reais para o chip 2. Para produzir cada chip 1, investe-se 3 homens-hora de trabalho especializado, 2 homens-hora de trabalho não especializado e 1 unidade de matéria-prima por semana. Para produzir cada chip 2, investe-se 4 homens-hora de trabalho especializado, 3 homens-hora de trabalho não especializado e 2 unidades de matéria-prima por semana. A empresa viabiliza 100 homens-hora de trabalho especializado, 70 homens-hora de trabalho não especializado e 30 unidades de matéria-prima semanais. O contrato de venda assinado obriga a produção semanal de pelo menos 3 unidades do chip 2 e qualquer quantidade fracionária é aceitável. Formular o problema para que a empresa obtenha lucro máximo.

5. **(Páginas 70 e 71 em [19])** A Viação Aérea Brasileira está estudando a compra de três tipos de aviões: Boeing 717 para as pontes aéreas de curta distância, Boeing 737-500 para vôos domésticos e internacionais de média distância e MD-11 para vôos internacionais de longa distância. Em um estudo preliminar, considerou-se que a capacidade máxima dos aviões a serem comprados será sempre preenchida para efeito de planejamento. Os dados de planejamento são: um avião do tipo Boeing 717 custa 5,1 milhões de dólares, tem uma receita teórica de 330 milhões de dólares e a empresa deve ter 30 pilotos aptos; um avião do tipo Boeing 737-500 custa 3,6 milhões de dólares, tem uma receita teórica de 300 milhões de dólares e a empresa deve ter 20 pilotos aptos; e um avião do tipo MD-11 custa 6,8 milhões de dólares, tem uma receita teórica de 420 milhões de dólares e a empresa deve ter 10 pilotos aptos. A verba disponível para as compras é de 220 milhões de dólares. Os pilotos de MD-11 podem pilotar todos os aviões da empresa, mas os demais pilotos só podem ser escalados às aeronaves a que foram habilitados. Cada aeronave necessita de dois pilotos para operar. As oficinas de manutenção podem acomodar até 40 Boeings 717. Um Boeing 737-500 equivale, em esforço de manutenção, a 3/4, e um MD-11 a 5/3, quando referidos ao Boeing 717. Formular um modelo de PL para o problema de otimizar as aquisições de aviões.

Capítulo 4

Poliedros e o Conjunto Viável

Aqui e nos próximos seis capítulos, trataremos do estudo dos fundamentos da PL. O que faremos neste caminho, então, será reescrever o que já existe na literatura, dando uma primeira olhada no conjunto viável como um poliedro e, em seguida, caracterizando-o como um conjunto poliedral com um número finito de pontos extremos e com pelo menos um ponto extremo quando não vazio, enunciando e demonstrando o Teorema Fundamental da PL, definindo o problema dual, enunciando e demonstrando o Lema de Farkas para, em seguida, enunciar e demonstrar o Teorema de Dualidade e, finalmente, definindo o problema primal-dual associando-o às condições de otimalidade para um PPL.

Iniciamos o nosso intuito com alguns resultados de convexidade. Nosso objetivo aqui é demonstrar que poliedros são conjuntos convexos e fechados e que o conjunto viável é um poliedro.

4.1 Poliedros

Iniciamos o nosso propósito definindo combinações lineares e, em seguida, definindo conjuntos convexos e demonstrando que a interseção finita de convexos é um conjunto convexo.

Definição 4.1.1 *Sejam dados q vetores* $x^1, x^2, \ldots, x^q \in R^n$.

(a) Dizemos que $x \in R^n$ *é uma combinação linear de* $x^1, x^2, \ldots, x^q \in R^n$*, quando existem q escalares* $\lambda_1, \lambda_2, \ldots, \lambda_q \in R$ *tais que*

$$x = \lambda_1 x^1 + \lambda_2 x^2 + \ldots + \lambda_q x^q;$$

(b) *dizemos que $x \in R^n$ é uma combinação linear afim, ou simplesmente combinação afim, de $x^1, x^2, \ldots, x^q \in R^n$, quando x é uma combinação linear e*

$$\lambda_1 + \lambda_2 + \ldots + \lambda_q = 1;$$

e

(c) *dizemos que $x \in R^n$ é uma combinação linear convexa, ou simplesmente combinação convexa, de $x^1, x^2, \ldots, x^q \in R^n$, quando x é uma combinação linear afim e*

$$\lambda_1, \lambda_2, \ldots, \lambda_q \in [0, 1].$$

Atenção: Que tal fazer o exercício 1 agora?

Definição 4.1.2 *Seja S um subconjunto de R^n.*

(a) *Dizemos que S é um conjunto afim, quando todas as combinações afins de quaisquer dois pontos de S pertencem a S; e*

(b) *dizemos que S é um conjunto convexo, quando todas as combinações convexas de quaisquer dois pontos de S pertencem a S.*

Atenção: Que tal fazer o exercício 2 agora?

Agora estamos prontos para demonstrar que a interseção (no nosso caso, finita) de conjuntos convexos é um conjunto convexo. A demonstração da proposição a seguir pode ser encontrada na página 90 em [27].

Proposição 4.1.3 *Suponha que $S_1, S_2, \ldots, S_q$ são subconjuntos convexos de R^n. Então,*

$$S = S_1 \cap S_2 \cap \ldots \cap S_q$$

é um conjunto convexo.

Demonstração: Se a interseção é vazia, então S é convexo porque não existe par de pontos no conjunto vazio. Caso contrário, fixemos arbitrariamente $x^1, x^2 \in S$. Tomemos $\lambda_1, \lambda_2 \in [0,1]$ com $\lambda_1 + \lambda_2 = 1$. Considere $i = 1, 2, \ldots, q$. Uma vez que $x^1, x^2 \in S$, segue-se pela definição de interseção de conjuntos que $x^1, x^2 \in S_i$, para cada i. Logo, para cada i, $x = \lambda_1 x^1 + \lambda_2 x^2 \in S_i$, pela convexidade de S_i. Segue-se, novamente pela definição de interseção de conjuntos, que $x \in S$. Pela arbitrariedade de $x^1, x^2 \in S$, S é um conjunto convexo. Isto finaliza a demonstração. ∎

Atenção: Que tal fazer o exercício 3 agora?

Neste ponto vamos definir um vetor que é fundamental para o estudo da PL.

Definição 4.1.4 *Seja S um subconjunto convexo de R^n. Um ponto x em S é denominado ponto extremo de S, quando x não é uma combinação linear convexa de quaisquer dois outros pontos distintos em S.*

Atenção: Que tal fazer o exercício 4 agora?

Agora vamos definir bolas: bola aberta e bola fechada. E, também, conjunto limitado.

Definição 4.1.5 *Considere um vetor $\bar{x} \in R^n$ e um número real positivo t.*

(a) A bola aberta de centro $\bar{x}$ e raio t é o conjunto dos pontos $x \in R^n$ cuja distância ao ponto $\bar{x}$ é menor do que t, isto é,

$$B(\bar{x}, t) = \{x \in R^n; \|x - \bar{x}\| < t\}.$$

(b) A bola fechada de centro $\bar{x}$ e raio t é o conjunto dos pontos $x \in R^n$ cuja distância ao ponto $\bar{x}$ é menor do que ou igual a t, isto é,

$$B[\bar{x}, t] = \{x \in R^n; \|x - \bar{x}\| \leq t\}.$$

(c) Dizemos que um conjunto S, $S \subset R^n$, é limitado, quando existe um escalar $\Gamma > 0$, tal que $\|x\| \leq \Gamma$ para todo $x \in S$.

Atenção: Que tal fazer o exercício 5 agora?

O próximo passo será definir conjunto fechado e demonstrar que a interseção (no nosso caso, finita) de conjuntos fechados é um conjunto fechado.

Definição 4.1.6 *Seja S um subconjunto de R^n.*

(a) Um ponto $\bar{x} \in R^n$ chama-se ponto de acumulação do conjunto S, quando toda bola aberta de centro $\bar{x}$ possui algum ponto de S, diferente de $\bar{x}$. Ou seja, para todo $\epsilon > 0$, deve existir $x \in S$ tal que $0 < \|x - \bar{x}\| < \epsilon$;

(b) um ponto $\bar{x} \in R^n$ diz-se aderente a um conjunto S, quando é limite de uma seqüência de pontos desse conjunto;

(c) o conjunto dos pontos aderentes a S chama-se o fecho de S, denotado por $fecho(S)$;

(d) um conjunto S chama-se fechado, quando possui todos os seus pontos aderentes, isto é, se $\lim x^k = \bar{x}$ e $x^k \in S$ para todo $k \in N$, então $\bar{x} \in S$; e

(e) um conjunto S chama-se compacto, quando S é limitado e fechado.

Atenção: Que tal fazer o exercício 6 agora?

A demonstração da proposição a seguir pode ser encontrada na página 40 em [33].

Proposição 4.1.7 *Suponha que $S_1, S_2, \ldots, S_q$ são subconjuntos fechados de R^n. Então,*

$$S = S_1 \cap S_2 \cap \ldots \cap S_q$$

é um conjunto fechado.

Demonstração: Se a interseção é vazia, então S é fechado porque o conjunto vazio possui zero pontos aderentes. Caso contrário, dado um ponto $\bar{x}$ aderente a S, devemos demonstrar que $\bar{x} \in S$. Com efeito, consideremos $i = 1, 2, \ldots, q$ e suponhamos $\bar{x}$ aderente a S. Por hipótese, existe uma seqüência (x^k), $x^k \in S$, para todo $k \in N$, tal que $\lim x^k = \bar{x}$. Para todo $k \in N$, uma vez que $x^k \in S$, segue-se pela definição de S, que $x^k \in S_i$, para todo i. Daí, $\bar{x}$ é aderente para cada S_i. Para todo i, uma vez que S_i é fechado segue-se que $\bar{x} \in S_i$. Pela definição de interseção de conjuntos, $\bar{x} \in S$. Isto finaliza a demonstração. ∎

Atenção: Que tal fazer o exercício 7 agora?

Afinal, o que são poliedros? A seguir, definimos poliedro como um poliedro convexo fechado, conforme [27].

Definição 4.1.8 *Sejam dados um vetor não nulo $a \in R^n$, denominado vetor normal, e um escalar $\delta \in R$.*

(a) O conjunto

$$H = \{x \in R^n; a^T x = \delta\}$$

é denominado um hiperplano;

(b) os conjuntos

$$H_l = \{x \in R^n; a^T x \leq \delta\}$$

e

$$H_u = \{x \in R^n; a^T x \geq \delta\}$$

são denominados semiespaços fechados; e

(c) um poliedro convexo fechado ou, simplesmente, poliedro, é um conjunto formado pela interseção de um número finito de semiespaços fechados.

Pela definição de poliedro, observamos que o conjunto vazio é um poliedro.

Atenção: Que tal fazer o exercício 8 agora?

A seguir definiremos um poliedro em particular, ou seja, um politopo.

Definição 4.1.9 *Seja S um subconjunto de R^n. Dizemos que um conjunto S é um politopo, quando S é um poliedro limitado.*

Atenção: Que tal fazer o exercício 9 agora?

Agora estamos prontos para demonstrar que um poliedro é um conjunto convexo e fechado.

Proposição 4.1.10 *Seja S um subconjunto de R^n. Se S é um poliedro, então S é um conjunto convexo e fechado.*

Demonstração: Suponha S um poliedro. Por definição, S é a interseção de um número finito de semiespaços fechados. Sem perda de generalidade, supomos que estes semiespaços fechados são da forma H_l. Assim, vamos demonstrar que o semiespaço fechado H_l é convexo. Com efeito, fixando arbitrariamente $x^1, x^2 \in H_l$ e tomando $\lambda_1, \lambda_2 \in [0, 1]$ tal que $\lambda_1 + \lambda_2 = 1$, segue-se pela definição de H_l, que

$$a^T(\lambda_1 x^1 + \lambda_2 x^2) = \lambda_1 a^T x^1 + \lambda_2 a^T x^2 \leq \lambda_1 \delta + \lambda_2 \delta = \delta.$$

Logo, pela arbitrariedade de $x^1, x^2 \in H_l$, concluímos que H_l é convexo. Agora, como S é o conjunto interseção de um número finito de conjuntos convexos da forma H_l, segue-se pela Proposição 4.1.3 que S é convexo. Finalmente, pela definição de S, uma vez que cada semiespaço fechado é fechado (veja página 39 em [6]), segue-se pela Proposição 4.1.6 que S é fechado. Isto finaliza a demonstração. ∎

Atenção: Que tal fazer o exercício 10 agora?

Finalmente, vamos demonstrar que o conjunto viável de um PPL é um poliedro.

Proposição 4.1.11 *Considere o PPL (P). O conjunto de soluções viáveis $\mathcal{X}$ é um poliedro, convexo e fechado.*

Demonstração: Usando a Proposição 4.1.10, basta-nos demonstrar que $\mathcal{X}$ é um poliedro. Por definição, $\mathcal{X}$ é o conjunto interseção de $2m+n$ semiespaços fechados. Portanto, $\mathcal{X}$ é um poliedro. Isto finaliza a demonstração. ∎

No próximo capítulo caracterizaremos definitivamente o conjunto viável de um PPL.

Atenção: Que tal fazer o exercício 11 agora?

4.2 Exercícios

1. Considere $x = [2,1]^T$. Encontre uma de suas combinações lineares, uma de suas combinações afins e uma de suas combinações convexas.

2. Sejam dados os subconjuntos do R^2, a saber:

$$S_1 = \emptyset,\ S_2 = \{(1,0)^T\},\ S_3 = \{x \in R^2; x_1 + x_2 = 2\}$$

   e

$$S_4 = \{x \in S_3; x_1, x_2 \geq 0\}.$$

   Representar graficamente os conjuntos que são:

   (a) conjuntos afins; e

   (b) conjuntos convexos.

3. Dê um contra-exemplo para demonstrar que a reunião finita de conjuntos convexos não é um conjunto convexo.

4. Desenhe conjuntos com zero, um, dois e uma infinidade de pontos extremos.

5. Dê um exemplo de uma bola aberta e de uma bola fechada na reta R.

6. Dê um exemplo de um conjunto fechado que não é um conjunto compacto e dê um outro exemplo de um conjunto compacto.

7. Sejam dados os subconjuntos do R^2, a saber:

$$S_1 = \{x \in R^2; x_1, x_2 \geq 0\} \text{ e } S_2 = \{x \in R^2; x_1 - x_2 = 1\}.$$

   Pede-se: representar graficamente os conjuntos fechados S_1, S_2 e $S_1 \cap S_2$.

8. Considere $a = [-1,-1]^T$ e $\delta = 2$. Represente graficamente os conjuntos H, H_l e H_u. Além disso, defina dois poliedros distintos e represente-os graficamente.

9. Dê um exemplo de um politopo e dê um exemplo de um poliedro que não é um politopo.

10. Todo hiperplano é um conjunto convexo e fechado? Demonstre ou dê um contra-exemplo.

11. Uma reta qualquer no R^2 é um conjunto poliedral, convexo e fechado. Intuitivamente falando, você acha que o conjunto viável de um PPL pode ser uma reta? Justifique.

Capítulo 5

Caracterização do Conjunto Viável

Continuamos com o nosso estudo dos fundamentos da PL. Nosso objetivo aqui é caracterizar o conjunto viável como um poliedro com pelo menos um ponto extremo quando não vazio.

5.1 Solução básica viável

Neste ponto, vamos reescrever a definição de ponto extremo de uma maneira, digamos, mais operacional.

Definição 5.1.1 *Sejam dados uma matriz A, $m \times n$, $0 < m < n$, e um vetor b em R^m. Considere um sistema de equações lineares $Ax = b$, tal que posto$(A) = m$.*

(a) Uma matriz quadrada B, $m \times m$, obtida de A, com m vetores coluna linearmente independentes denomina-se matriz base de A ou base de A. Uma matriz N, $m \times (n - m)$, obtida de A, com os $n - m$ vetores coluna restantes denomina-se matriz não base.

(b) Considere uma matriz base B, $m \times m$. O conjunto de índices correspondentes a esta matriz base B, no sistema $Ax = b$, chama-se conjunto de índices base. O conjunto com os demais $n - m$ índices chama-se conjunto de índices não base. Denotamos o conjunto de índices base por I_B e o conjunto de índices não base por I_N.

(c) *Considere uma matriz base B, $m \times m$. As variáveis correspondentes a esta matriz base B, no sistema $Ax = b$, chamam-se variáveis básicas. As demais $n - m$ variáveis chamam-se variáveis não básicas. Denotamos o vetor de variáveis básicas por x_B e o vetor de variáveis não básicas por x_N.*

(d) *Anulando as $n - m$ variáveis não básicas, obtemos um sistema compatível determinado, constituído de m equações e m incógnitas. Determinando o valor das variáveis básicas, obtemos uma solução básica. Ou seja, $x \in R^n$ é uma solução básica, quando $x_N = 0$ e x_B é a solução do sistema linear $Bx_B = b$.*

(e) *Uma solução básica onde as variáveis básicas são não negativas denomina-se solução básica viável.*

(f) *Uma solução básica viável onde existe ao menos uma variável básica nula denomina-se solução básica viável degenerada.*

Por conveniência, suponhamos que $x_1, \dots, x_m$ são as variáveis básicas, coordenadas de x_B, e $x_{m+1}, \dots, x_n$ são as variáveis não básicas, coordenadas de x_N. Para a resolução do sistema $Ax = b$ buscamos exprimir x_B em função de x_N, a saber:

$$Ax = b \Rightarrow Bx_B + Nx_N = b \Rightarrow x_B = B^{-1}b - B^{-1}Nx_N.$$

Para $x_N = 0$, $x_B = B^{-1}b$, e x é uma solução básica. Se $x_B = B^{-1}b \geq 0$, para $x_N = 0$, então x é uma solução básica viável.

Atenção: Que tal fazer os exercícios 1, 2 e 3 agora?

Se demonstrarmos que para um PPL um ponto extremo é uma solução básica viável e vice-versa, então obteremos uma caracterização para pontos extremos mais operacional. É o que faremos no próximo teorema, cuja demonstração pode ser encontrada nas páginas 31, 32 e 33 em [2].

Teorema 5.1.2 *Considere o PPL (P). Um ponto viável $x \in \mathcal{X}$ é ponto extremo se, e somente se, x é uma solução básica viável.*

Demonstração:

($\Longrightarrow$) Aqui vamos demonstrar que se x é um ponto extremo então x é uma solução básica viável (*sbv*). Suponhamos, então, que x é um ponto extremo de $\mathcal{X}$. Se x é o vetor nulo, então como a matriz A tem posto completo, segue-se que x é uma *sbv* para alguma matriz base de A. Sem perda de generalidade, suponhamos que as primeiras q componentes de x são positivas. Pela viabilidade de x, para $j = 1, \ldots, q$, $x_j > 0$ e $A_1x_1 + \ldots + A_qx_q = b$, onde A_j é a j-ésima coluna da matriz A. Para demonstrar que x é uma *sbv*, devemos mostrar que os vetores $A_1, \ldots, A_q$, associados às componentes positivas de x, são linearmente independentes (*li*). Suponhamos, por contradição, que estes vetores são linearmente dependentes, isto é, existem números λ_j, $j = 1, \ldots, q$, não todos nulos, tais que

$$\sum_{j=1}^{q} A_j\lambda_j = 0. \tag{5.1}$$

Selecionando

$$\sigma = \min\{\frac{x_j}{|\lambda_j|}; \lambda_j \neq 0,\ j = 1, \ldots, q\},$$

podemos escolher ϵ, $0 < \epsilon < \sigma$, tal que $x_j + \epsilon\lambda_j > 0$ e $x_j - \epsilon\lambda_j > 0$, para $j = 1, \ldots, q$. Assim, definimos x^1 e x^2 por

$$x^1 = x + \epsilon\lambda \geq 0 \text{ e } x^2 = x - \epsilon\lambda \geq 0, \tag{5.2}$$

onde $\lambda = [\lambda_1, \ldots, \lambda_q, 0 \ldots, 0]^T$. Segue-se pela definição de λ e por (5.1), que $A\lambda = 0$. Logo, pela linearidade de A, $Ax^1 = b$ e $Ax^2 = b$, de tal maneira que, usando (5.2), $x^1, x^2 \in \mathcal{X}$ e tanto x^1 quanto x^2 são diferentes de x. Além disso, $x = (x^1 + x^2)/2$, isto é, x é uma combinação convexa de dois outros pontos distintos, contradizendo o fato de que x é um ponto extremo. Logo, $A_1, \ldots, A_q$ são vetores *li* e, portanto, se x é um ponto extremo de $\mathcal{X}$, x é uma *sbv* de $\mathcal{X}$. Isto finaliza a primeira parte da demonstração.

($\Longleftarrow$) Aqui vamos demonstrar que se x é uma *sbv* então x é um ponto extremo. Suponhamos, então, que x é uma *sbv* de $\mathcal{X}$. Se necessário, rearranjando suas componentes,

$$x = [x_1, \ldots, x_m, 0, \ldots, 0]^T, \tag{5.3}$$

uma vez que x é uma *sbv*. Além disso, pela viabilidade de x, $x \geq 0$ e

$$Ax = Bx_B = b, \tag{5.4}$$

onde B é a matriz base, obtida das m primeiras colunas de A. Se

$$x = (x^1 + x^2)/2 \in \mathcal{X}$$

para dois pontos quaisquer $x^1, x^2 \in \mathcal{X}$, segue-se por (5.3) que as m primeiras componentes de x^1 e x^2 são não negativas e as $n-m$ componentes restantes são iguais a zero. Pelo fato de B ser uma matriz base, o sistema (5.4) possui uma única solução, logo, $x = x^1 = x^2$. Portanto, x é um ponto extremo de $\mathcal{X}$. Isto finaliza a demonstração. ∎

Assim, podemos calcular pontos extremos através do cálculo de soluções básicas viáveis. Além disso, devemos observar que a correspondência entre pontos extremos e soluções básicas viáveis não é em geral um-a-um.

Atenção: Que tal fazer o exercício 4 agora?

O número de pontos extremos em qualquer conjunto viável de um PPL é finito? A resposta é sim, e demonstraremos isto agora mesmo. Esta demonstração pode ser encontrada na página 107 em [11].

Corolário 5.1.3 *Considere o PPL (P). O conjunto de soluções viáveis tem um número finito de pontos extremos.*

Demonstração: Uma vez que $posto(A) = m < n$, dos n vetores coluna de A, existem no máximo

$$\begin{pmatrix} n \\ m \end{pmatrix} = \frac{n!}{m!(n-m)!}$$

conjuntos de m vetores linearmente independentes, o que significa no máximo combinações de n tomados m a m soluções básicas. Por definição, o número de soluções básicas viáveis é menor do que ou igual ao número de soluções básicas. Logo, usando o teorema anterior, segue-se que existe um número finito de pontos extremos do conjunto viável $\mathcal{X}$, finalizando a demonstração. ∎

Atenção: Que tal fazer o exercício 5 agora?

Na próxima seção formalizaremos a geomeria do conjunto viável de um PPL.

5.2 A caracterização de $\mathcal{X}$

Nesta seção, caracterizamos o conjunto viável $\mathcal{X}$ de um PPL através da formalização de sua geometria: $\mathcal{X}$ é um poliedro com um número finito de pontos extremos e, quando não vazio, possui ao menos um ponto extremo.

O próximo resultado lança as bases para a demonstração de existência de pontos extremos em um conjunto viável de um PPL. A demonstração deste teorema pode ser encontrada nas páginas 108 e 109 em [11].

Teorema 5.2.1 *Considere o PPL (P). Todo ponto viável x pode ser escrito como uma combinação linear convexa*

$$x = \beta\hat{x} + (1-\beta)\tilde{x}, \ \ 0 < \beta \leq 1,$$

onde $\hat{x}$ é ponto extremo de $\mathcal{X}$ e $\tilde{x} \in \mathcal{X}$.

Demonstração: Sem perda de generalidade, seja $x = (x_1, \ldots, x_q, 0 \ldots, 0)^T$, um ponto de $\mathcal{X}$, cujas q primeiras componentes são positivas. Demonstraremos o teorema usando indução finita em q.

Para $q = 0$: como $posto(A) = m$, podemos sempre encontrar uma base à qual seja possível associar a solução básica viável x, ou seja, pelo Teorema 5.1.2, o ponto extremo x. Temos, então, $x = \hat{x}$ com $\beta = 1$.

Para $q > 0$: vamos admitir que já tenhamos demonstrado o teorema para o caso em que x possui no máximo $q-1$ componentes positivas. Demonstremos agora que podemos estender esta afirmação para o caso em que x possua q componentes positivas. Consideremos dois casos.

(i) $A_1, \ldots, A_q$ são linearmente independentes: neste caso necessariamente $q \leq m$ e podemos, então, se necessário, adicionar às colunas $A_1, \ldots, A_q$, $m-q$ colunas de A formando uma base. Para esta base podemos associar a solução básica viável x, ou seja, pelo Teorema 5.1.2, o ponto extremo x. Temos, então, novamente $x = \hat{x}$ com $\beta = 1$.

(ii) $A_1, \ldots, A_q$ são linearmente dependentes: por definição, existem números $\lambda_1, \ldots, \lambda_q$, não todos nulos, tais que

$$\sum_{j=1}^{q} A_j \lambda_j = 0.$$

Assim, como na demonstração do Teorema 5.1.2, podemos definir

$$x = \frac{x^1 + x^2}{2}, \tag{5.5}$$

onde $x^1 = x + \sigma\lambda \in \mathcal{X}$ e $x^2 = x - \sigma\lambda \in \mathcal{X}$, onde

$$\sigma = \min\{\frac{x_j}{|\lambda_j|}; \lambda_j \neq 0,\ j = 1, \ldots, q\} = \frac{x_k}{|\lambda_k|}$$

e

$$\lambda = [\lambda_1, \ldots, \lambda_q, 0 \ldots, 0]^T.$$

Digamos que $\lambda_k < 0$, isto é, $\sigma = -x_k/\lambda_k$. A outra situação, isto é, $\lambda_k > 0$, é análoga. Temos, então, para a k-ésima componente de x^1,

$$x_k^1 = x_k + \sigma\lambda_k = x_k - \frac{x_k}{\lambda_k}\lambda_k = 0.$$

Como

$$x_k^1 = x_{q+1}^1 = \ldots = x_n^1 = 0,$$

x^1 tem no máximo $q - 1$ componentes positivas. De acordo com a hipótese de indução, temos,

$$x^1 = \alpha\hat{x} + (1 - \alpha)\bar{x},\ 0 < \alpha \leq 1, \tag{5.6}$$

onde $\hat{x}$ é ponto extremo de $\mathcal{X}$ e $\bar{x} \in \mathcal{X}$. Substituindo (5.5) em (5.6),

$$\begin{aligned} x &= \frac{\alpha\hat{x} + (1-\alpha)\bar{x} + x^2}{2} \\ &= \frac{\alpha}{2}\hat{x} + \frac{2-\alpha}{2}(\frac{1-\alpha}{2-\alpha}\bar{x} + \frac{1}{2-\alpha}x^2). \end{aligned}$$

Fazendo

$$\tilde{x} = \frac{1-\alpha}{2-\alpha}\bar{x} + \frac{1}{2-\alpha}x^2,$$

temos que $\tilde{x} \in \mathcal{X}$, porque $\mathcal{X}$ é convexo e $\tilde{x}$ é combinação linear convexa de dois pontos viáveis $\bar{x}$ e x^2, uma vez que

$$\frac{1-\alpha}{2-\alpha} + \frac{1}{2-\alpha} = 1 \quad \text{e} \quad \frac{1-\alpha}{2-\alpha}, \frac{1}{2-\alpha} \in [0,1].$$

Finalmente, fazendo $\beta = \alpha/2$, podemos, então, escrever

$$x = \beta\hat{x} + (1-\beta)\tilde{x}, \ 0 < \beta \leq 1,$$

onde $\hat{x}$ é ponto extremo de $\mathcal{X}$ e $\tilde{x} \in \mathcal{X}$.

Assim, supondo a hipótese verdadeira para $q-1$ componentes positivas foi possível demonstrá-la para q componentes positivas. Como a hipótese é verdadeira para $q = 0$, finalizamos a demonstração. ∎

Atenção: Que tal fazer o exercício 6 agora?

O próximo resultado caracteriza o conjunto viável de um PPL, formalizando assim, a sua geometria.

Corolário 5.2.2 *Considere o PPL (P). O conjunto viável $\mathcal{X}$ é um poliedro e, quando não vazio, possui ao menos um ponto extremo. Além disso, o número de pontos extremos é finito.*

Demonstração: Pela Proposição 4.1.11, $\mathcal{X}$ é um poliedro, pelo Corolário 5.1.3, o número de pontos extremos é finito e, usando o teorema anterior, $\mathcal{X}$ possui ao menos um ponto extremo, quando não vazio. Isto finaliza a demonstração. ∎

No próximo capítulo estudaremos o conjunto de soluções ótimas de um PPL. Além disso, enunciaremos e demonstraremos o Teorema Fundamental da PL.

Atenção: Que tal fazer o exercício 7 agora?

5.3 Exercícios

1. Considere

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} \text{ e } b = \begin{bmatrix} 3 \\ 1 \end{bmatrix}.$$

   Desenvolva $Ax = b$ objetivando isolar x_B, para alguma base B.

2. Considere o sistema de desigualdades

$$\begin{array}{ll} x_1 & \leq 2 \\ \quad\quad x_2 & \leq 3 \\ x_1 + \quad x_2 & \leq 3 \\ 3x_1 + 3x_2 & \leq 9 \\ x_1 \geq 0, \quad x_2 \geq 0. & \end{array}$$

   Transforme as desigualdades "menor do que ou igual a" em igualdades. A partir daí, defina uma matriz A e um vetor $b \in R^4$. Defina uma matriz base B associada à matriz A. Pede-se, para a base que você escolheu, se existir:

   (a) o conjunto de índices base e o conjunto de índices não base;

   (b) as variáveis básicas e as variáveis não básicas;

   (c) a solução básica;

   (d) a solução básica viável; e

   (e) a solução básica viável degenerada.

3. Desenhe o conjunto viável no R^2 do exercício anterior e identifique as sete soluções básicas viáveis degeneradas. Tire conclusões sobre o seu desenho. Em seguida, desenhe o conjunto viável em R^3 para o PPL

$$\begin{array}{lll} min & -x_1 & \\ s.\ a: & x_1 + x_2 + x_3 + x_4 & = 4 \\ & 0,8x_1 + x_2 + 2x_3 + x_5 & = 4 \\ & x_1, \ x_2, \ x_3, \ x_4, \ x_5 & \geq 0, \end{array}$$

   e identifique as quatro soluções básicas viáveis degeneradas. Refute a conclusão de que sempre existem restrições redundantes em uma solução básica viável degenerada. Observe que neste último problema de PL, $n - m = 3 > 2$ (veja página 86 em [21]).

4. Considere o PPL

$$\begin{array}{lll} minimizar & x_1 - x_2 & \\ sujeito\ a: & x_1 + x_2 & \leq 3 \\ & x_1 & \leq 2 \\ & \quad\; x_2 & \leq 3 \\ & x_1, \quad x_2 \geq 0. & \end{array}$$

   Pede-se:

   (a) coloque o PPL no formato padrão;
   (b) verifique que todos os pontos extremos são soluções básicas viáveis e vice-versa; e
   (c) verifique que o número de pontos extremos é menor do que ou igual ao número de soluções básicas viáveis. Por quê?

5. Considere o PPL do exercício anterior. Pede-se:

   (a) quais são as soluções básicas e as soluções básicas viáveis; e
   (b) quantas são as soluções básicas e as soluções básicas viáveis.

6. Considere o PPL do exercício anterior. Represente os pontos abaixo como combinação linear convexa de pelo menos um ponto extremo e algum outro ponto viável.

   (a) $x = [1, 1]^T$;
   (b) $x = [0, 0]^T$; e
   (c) $x = [1, 0]^T$.

7. Encontre um poliedro, convexo, fechado e não vazio, tal que um conjunto viável de qualquer PPL não pode assumir. Por quê?

Capítulo 6

Caracterização do Conjunto de Soluções Ótimas

Continuamos com o nosso estudo dos fundamentos da PL. Nosso objetivo aqui é caracterizar o conjunto de soluções ótimas como um poliedro que, quando não vazio, possui uma única ou uma infinidade de soluções ótimas. Quando o conjunto de soluções ótimas é vazio, vimos no Capítulo 2 que o problema de PL pode ser ilimitado ou inviável.

Inicialmente, enunciamos e demonstramos o Teorema Fundamental da PL para, em seguida, formalizar a geometria do conjunto de soluções ótimas.

6.1 O Teorema Fundamental da PL

Os dois últimos capítulos nos auxiliam para a demonstração e para o entendimento geométrico deste teorema. Assim, estamos prontos para enunciar e demonstrar o Teorema Fundamental da PL, cuja demonstração pode ser encontrada na página 111 em [11].

Teorema 6.1.1 *Considere o PPL* (P)*. Se* (P) *admite solução ótima, então uma solução ótima é atingida em ao menos um ponto extremo do conjunto viável.*

Demonstração: Seja x^* uma solução ótima do PPL (P). Pelo Teorema 5.2.1, podemos escrever

$$x^* = \beta\hat{x} + (1-\beta)\tilde{x}, 0 < \beta \leq 1,$$

onde $\hat{x}$ é um ponto extremo de $\mathcal{X}$ e $\tilde{x} \in \mathcal{X}$. Pela linearidade da função objetivo,

$$c^T x^* = c^T(\beta\hat{x} + (1-\beta)\tilde{x}) = \beta c^T\hat{x} + (1-\beta)c^T\tilde{x}.$$

Como x^* é uma solução ótima, $c^T x^* \leq c^T\tilde{x}$. Substituindo esta desigualdade na última igualdade,

$$c^T x^* \geq \beta c^T\hat{x} + (1-\beta)c^T x^*.$$

Daí resulta $\beta c^T x^* \geq \beta c^T\hat{x}$, ou seja, como $\beta > 0$,

$$c^T x^* \geq c^T\hat{x}.$$

Mas como x^* é uma solução ótima,

$$c^T x^* \leq c^T\hat{x}.$$

Das duas últimas desigualdades resulta $c^T x^* = c^T\hat{x}$. Evidentemente, podemos ter $x^* = \hat{x}$, isto é, x^* é ponto extremo. Caso tenhamos, no entanto, $x^* \neq \hat{x}$, existe mais de uma solução ótima entre as quais se encontra ao menos um ponto extremo, isto é $\hat{x}$. Isto finaliza a demonstração. ∎

Atenção: Que tal fazer o exercício 1 agora?

Na próxima seção formalizaremos a geometria do conjunto de soluções ótimas de um PPL.

6.2 A caracterização de $\mathcal{X}(P)$

Nesta seção, caracterizamos o conjunto de soluções ótimas $\mathcal{X}(P)$ através da formalização de sua geometria: quando não vazio, $\mathcal{X}(P)$ é um poliedro que possui uma única ou uma infinidade de soluções ótimas.

Iniciamos o nosso propósito com o seguinte resultado.

Teorema 6.2.1 *Considere o PPL (P). Cada combinação linear convexa de soluções ótimas é por sua vez também uma solução ótima.*

Demonstração: Sejam $x^1, \ldots, x^q$ soluções ótimas do PPL (P). Seja o ponto $x^* \in R^n$ uma combinação linear convexa qualquer de x^k, $k = 1, \ldots, q$. Isto é,

$$x^* = \sum_{k=1}^{q} \alpha_k x^k, \ \sum_{k=1}^{q} \alpha_k = 1, \ \alpha_k \geq 0, \ k = 1, \ldots, q.$$

Então, $x^* \geq 0$ e, usando a linearidade de A e a viabilidade de $x^1, \ldots, x^q$, $Ax^* = b$. Então, $x^* \in \mathcal{X}$. Por outro lado,

$$c^T x^* = c^T (\sum_{k=1}^{q} \alpha_k x^k) = \alpha_1 c^T x^1 + \ldots + \alpha_q c^T x^q.$$

Uma vez que $x^1, \ldots, x^q$ são soluções ótimas, segue-se que

$$c^T x^1 = \ldots = c^T x^q = v(P).$$

Então, usando o fato de que $\sum_{k=1}^{q} \alpha_k = 1$, obtemos

$$c^T x^* = \sum_{k=1}^{q} \alpha_k v(P) = v(P).$$

Portanto, $x^* \in \mathcal{X}(P)$ e pela arbitrariedade de x^* concluímos a demonstração.
∎

Atenção: Que tal fazer o exercício 2 agora?

Corolário 6.2.2 *Considere o PPL (P). Se (P) possui mais de uma solução ótima, então possui uma infinidade de soluções ótimas.*

Demonstração: Considere $x^1, x^2 \in \mathcal{X}(P)$ tais que $x^1 \neq x^2$. Segue-se pelo teorema anterior que

$$x = \alpha x^1 + (1 - \alpha) x^2 \in \mathcal{X}(P),$$

para todo $\alpha \in [0, 1]$. Portanto, $\mathcal{X}(P)$ possui uma infinidade de soluções ótimas, finalizando a demonstração. ∎

Atenção: Que tal fazer o exercício 3 agora?

Finalmente, vamos caracterizar o conjunto de soluções ótimas $\mathcal{X}(P)$.

Teorema 6.2.3 *Considere o PPL (P). O conjunto de soluções ótimas é um poliedro e, quando não vazio, possui uma única ou uma infinidade de soluções ótimas.*

Demonstração: Por definição, $\mathcal{X}(P)$ é o conjunto interseção de um número finito, $2m+n+2$, semiespaços fechados. Portanto, $\mathcal{X}(P)$ é um poliedro. Pelo corolário anterior, $\mathcal{X}(P)$ possui uma única ou uma infinidade de soluções ótimas quando não vazio. Isto finaliza a demonstração. ∎

No próximo capítulo iniciaremos o estudo de dualidade em Programação Linear.

Atenção: Que tal fazer os exercícios 4 e 5 agora?

6.3 Exercícios

1. Dê um exemplo de um PPL com uma solução ótima. Verifique que pelo menos uma solução ótima é um ponto extremo.

2. Seja o PPL no formato padrão definido por

$$A = [1\ 1],\ b = 2 \text{ e } c = [2, 2]^T.$$

   Encontrar todos os pontos extremos que são soluções ótimas e verificar que toda combinação linear convexa destes minimizadores encontrados é um ponto de mínimo também.

3. Dê um exemplo para um PPL:
   (a) com uma única solução ótima;
   (b) com uma infinidade de soluções ótimas;
   (c) ilimitado; e
   (d) inviável.

4. Considere os problemas de PL do exercício anterior. Verifique que os conjuntos de soluções ótimas são poliedros.

5. Demonstre que o conjunto de soluções ótimas é um conjunto convexo.

Capítulo 7

O Problema Dual

Continuamos com o nosso estudo dos fundamentos da PL. Nosso objetivo aqui é definir o problema de Programação Linear dual, para o PPL primal (P) no formato padrão.

Iniciamos o nosso intuito estudando um pouco de Álgebra Linear.

7.1 Preliminares

Inicialmente, vamos definir ortogonalidade entre dois subespaços vetoriais.

Definição 7.1.1 *Considere o espaço vetorial R^n. Dois subespaços V e W de R^n são ortogonais, quando qualquer vetor $v \in V$ é ortogonal a qualquer vetor $w \in W$, isto é,*

$$v^T w = 0, \text{ para todo } v \in V \text{ e para todo } w \in W.$$

Atenção: Que tal fazer o exercício 1 agora?

Considere a transformação linear definida pela matriz $A \in R^{m \times n}$. Dois subespaços importantes do espaço vetorial R^n estão associados com esta transformação: o espaço nulo de A, definido por

$$\mathcal{N}(A) = \{x \in R^n; Ax = 0\}$$

e seu complemento ortogonal, o espaço linha de A (também chamado espaço imagem de A^T ou espaço coluna de A^T), definido por

$$\mathcal{R}(A^T) = \{x \in R^n; x = A^T z, z \in R^m\}.$$

A proposição a seguir relaciona o espaço nulo de A e o espaço linha de A, através da definição de ortogonalidade entre dois subespaços. A demonstração pode ser encontrada nas páginas 136 e 137 em [55].

Proposição 7.1.2 *Seja o espaço R^n. O subespaço nulo de A é ortogonal ao subespaço linha de A.*

Demonstração: Seja o espaço R^n. Suponhamos $u \in \mathcal{N}(A)$ e $v \in \mathcal{R}(A^T)$. Por definição, $Au = 0$ e $v = A^T z$, para algum $z \in R^m$. Daí, usando propriedades de transposição de matrizes,

$$v^T u = (A^T z)^T u = z^T (A^T)^T u = z^T A u = z^T 0 = 0.$$

Pela definição de ortogonalidade, concluímos a demonstração. ∎

Atenção: Que tal fazer o exercício 2 agora?

Desta proposição, qualquer vetor $d \in R^n$ pode ser unicamente decomposto (soma direta) como $d = d_p + d_{\tilde{p}}$, onde $d_p \in \mathcal{N}(A)$ e $d_{\tilde{p}} \in \mathcal{R}(A^T)$. Os vetores d_p e $d_{\tilde{p}}$ são, respectivamente, a projeção de d no espaço nulo de A e no seu complemento ortogonal. A projeção de d no espaço nulo de A é o ponto no $\mathcal{N}(A)$ com a menor distância euclidiana para d. Esta é a definição mais usual de projeção:

$$d_p = argmin\{\|x - d\|; x \in \mathcal{N}(A)\}.$$

A demonstração da proposição a seguir pode ser encontrada na página 9 em [23].

Proposição 7.1.3 *Seja $A \in R^{m \times n}$ uma matriz de posto igual a m, $m \leq n$. Então, a matriz AA^T é não singular.*

Demonstração: Suponha por contradição que $AA^T d = 0$, para algum $d \neq 0$. Multiplicando d^T à esquerda desta igualdade,

$$d^T A A^T d = (A^T d)^T (A^T d) = 0.$$

Esta última igualdade é equivalente a $\|A^T d\|^2 = 0$, ou equivalentemente, $A^T d = 0$. Assim, as colunas de A^T são linearmente dependentes, isto é, as linhas de A são linearmente dependentes, contrariando o fato do posto de A ser igual a m. Isto finaliza a demonstração. ∎

Atenção: Que tal fazer o exercício 3 agora?

Uma vez que o operador projeção é linear (veja página 116 em [27]), podemos representá-lo por uma matriz. Isto é o que afirma a próxima proposição, cuja demonstração pode ser encontrada na página 9 em [23].

Proposição 7.1.4 *Sejam $A \in R^{m\times n}$ uma matriz de posto igual a m, $m \leq n$ e $d \in R^n$ um vetor arbitrário. Então,*

$$d_p = P_A d \quad e \quad d_{\tilde{p}} = \tilde{P}_A d,$$

onde

$$P_A = I - A^T(AA^T)^{-1}A \quad e \quad \tilde{P}_A = I - P_A.$$

Demonstração: Sabemos que d pode ser decomposto em $d = d_p + d_{\tilde{p}}$. Como $d_{\tilde{p}} \in \mathcal{R}(A^T)$, existe $z \in R^m$ tal que $d_{\tilde{p}} = A^T z$. Logo, $d = d_p + A^T z$. Multiplicando esta igualdade por A, $Ad = Ad_p + AA^T z$. Como $Ad_p = 0$ por definição, $AA^T z = Ad$. E, usando a proposição anterior, como AA^T é não singular, $z = (AA^T)^{-1}Ad$. Agora, substituindo esta igualdade em z na expressão $d_{\tilde{p}} = A^T z$ e, em seguida, substituindo $d_{\tilde{p}}$ em $d_p = d - d_{\tilde{p}}$, obtemos respectivamente,

$$d_{\tilde{p}} = A^T(AA^T)^{-1}Ad \;\text{ e }\; d_p = (I - A^T(AA^T)^{-1}A)d.$$

Tomando $P_A = I - A^T(AA^T)^{-1}A$ e $\tilde{P}_A = I - P_A$, concluímos que $d_p = P_A d$ e $d_{\tilde{p}} = \tilde{P}_A d$, finalizando a demonstração. ∎

Atenção: Que tal fazer o exercício 4 agora?

Agora vamos demonstrar que a matriz de projeção no espaço nulo de A, denotada por P_A, e a matriz de projeção no espaço linha de A, denotada por $\tilde{P}_A$ são matrizes simétricas e idempotentes. Esta demonstração pode ser encontrada nas páginas 158 e 159 em [55].

Proposição 7.1.5 *A matriz de projeção P_A é uma matriz simétrica, isto é, $P_A^T = P_A$; e uma matriz idempotente, isto é, $P_A^2 = P_A$. O mesmo vale para a matriz de projeção $\tilde{P}_A$.*

Demonstração: Temos pela proposição anterior que

$$P_A = I - A^T(AA^T)^{-1}A \;\text{ e }\; \tilde{P}_A = I - P_A.$$

Usando as propriedades

$$((AA^T)^{-1})^T = ((AA^T)^T)^{-1} \text{ e } (AA^T)^T = (A^T)^T A^T = AA^T,$$

segue-se por transposição em P_A que $P_A^T = P_A$. Agora, usando a simetria de P_A, concluímos que $\tilde{P}_A = I - P_A$ também é uma matriz simétrica.

Finalmente, desenvolvendo a igualdade

$$P_A^2 = (I - A^T(AA^T)^{-1}A)(I - A^T(AA^T)^{-1}A),$$

obtemos

$$P_A^2 = P_A - A^T(AA^T)^{-1}A + A^T(AA^T)^{-1}AA^T(AA^T)^{-1}A = P_A.$$

Agora, usando a idempotência de P_A, concluímos que $\tilde{P}_A = I - P_A$ também é uma matriz idempotente. Isto completa a demonstração. ∎

Por curiosidade, vale a recíproca desta proposição (veja páginas 158 e 159 em [55]).

Atenção: Que tal fazer o exercício 5 agora?

Neste ponto, estamos prontos para iniciarmos o estudo do problema dual através de sua definição, que é o nosso objetivo para este capítulo.

7.2 O problema dual

O problema de Programação Linear denominado o dual do problema (P), é o seguinte problema de Otimização:

$$\begin{array}{lll} (D) & \textit{maximizar} & b^T y \\ & \textit{sujeito a:} & A^T y + s = c \\ & & s \geq 0, \end{array}$$

onde o vetor $y \in R^m$ é denominado variável dual e incluímos explicitamente um vetor com componentes não negativas $s \in R^n$ denominado folga dual.

Considere os problemas (P) e (D). Estes problemas são denominados, respectivamente, problema primal e problema dual, de acordo com o seguinte teorema, cuja demonstração pode ser encontrada na página 97 em [21].

Teorema 7.2.1 *O dual de um problema dual é o primal.*

Demonstração: Considere o problema dual (D). Usando a Proposição 2.1.2 e trocando sinais, o problema (D) é equivalente ao problema

$$\begin{array}{ll} \textit{minimizar} & -b^T y \\ \textit{sujeito a:} & -A^T y - s = -c \\ & s \geq 0. \end{array}$$

Definindo $y = \bar{y} - \hat{y}$, $\bar{y} \geq 0$ e $\hat{y} \geq 0$, e substituindo neste último problema, então seu dual sem acrescentar as variáveis de folga é o problema

$$\begin{array}{lrl} \textit{maximizar} & -c^T x & \\ \textit{sujeito a:} & -(A^T)^T x & \leq -(b^T)^T \\ & (A^T)^T x & \leq (b^T)^T \\ & -x & \leq 0. \end{array}$$

Este último problema é equivalente ao problema primal (P). Isto finaliza a demonstração. ∎

Observe, conforme Capítulo 3, por exemplo, que existem várias formulações para um PPL. Todavia, o romantismo está no fato de que para cada PPL existe o seu par: o seu dual. Isto se dá pelo teorema anterior e porque vimos no Capítulo 2 que qualquer PPL pode ser colocado no formato padrão.

Atenção: Que tal fazer o exercício 6 agora?

Seguem-se algumas definições associadas ao problema (D). A função linear $y \mapsto b^T y$ é chamada função objetivo dual ou simplesmente função objetivo. Similarmente, temos: o conjunto

$$\mathcal{S} = \{(y, s) \in R^m \times R^n; A^T y + s = c, s \geq 0\},$$

é chamado conjunto viável e um ponto $(y, s) \in \mathcal{S}$ é denominado ponto viável. O conjunto $\{(y, s) \in R^m \times R^n; s > 0\}$ é chamado conjunto de pontos interiores e um ponto deste conjunto é denominado ponto interior. O conjunto

$$\mathcal{S}^0 = \{(y, s) \in \mathcal{S}; s > 0\}$$

é chamado conjunto de pontos interiores viáveis e um ponto $(y, s) \in \mathcal{S}^0$ é denominado ponto interior viável. O número $v(D) = \max\{b^T y; (y, s) \in \mathcal{S}\}$, quando existe, é denominado o valor ótimo ou custo ótimo. O conjunto

$$\mathcal{S}(D) = \{(y, s) \in \mathcal{S}; b^T y = v(D)\}$$

é chamado conjunto de soluções ótimas e um ponto $(y, s) \in \mathcal{S}(D)$ é denominado solução ótima ou maximizador ou ponto de máximo.

Atenção: Que tal fazer o exercício 7 agora?

O próximo resultado fornece uma caracterização para uma folga dual viável para um PPL dual, devido a Gonzaga (veja Lema 3.1 em [24]).

Lema 7.2.2 *$s \in R^n$ é uma folga dual viável para o problema (D) se, e somente se,*

(a) $s \geq 0$ e $P_A s = P_A c$; ou equivalentemente,

(b) $s \geq 0$ e $s = P_A c - \gamma$, com $\gamma \perp \mathcal{N}(A)$.

Demonstração: A equivalência entre (a) e (b) se segue pela definição de projeção, isto é, $s \geq 0$ e, usando a Proposição 7.1.2 e a Proposição 7.1.5,

$$P_A s = P_A(P_A c - \gamma) = P_A P_A c - P_A \gamma = P_A c.$$

Do problema dual (D), s é uma folga dual viável se, e somente se, $s \geq 0$ e $c - s = A^T y$ para algum $y \in R^m$. Esta última condição é equivalente a $s \geq 0$ e $c - s$ ortogonal a $\mathcal{N}(A)$, uma vez que o espaço nulo de A e o espaço linha de A são complementarmente ortogonais. Mas, isto é equivalente a $s \geq 0$ e $P_A(c-s) = P_A c - P_A s = 0$, estabelecendo (a). Isto completa a demonstração. ∎

Atenção: Que tal fazer o exercício 8 agora?

Com a suposição sobre o posto de A, existe $\tilde{x}$ tal que $A\tilde{x} = b$. Então, para qualquer $(y, s) \in \mathcal{S}$, isto é, no conjunto viável dual,

$$b^T y = (A\tilde{x})^T y = \tilde{x}^T (A^T y) = \tilde{x}^T c - \tilde{x}^T s.$$

Assim, usando a Proposição 2.1.2,

$$\max\{b^T y\} = \max\{\tilde{x}^T c - \tilde{x}^T s\} = \tilde{x}^T c + \max\{-\tilde{x}^T s\} = \tilde{x}^T c - \min\{\tilde{x}^T s\}.$$

Daí, usando o lema anterior, segue-se que o problema (D) é equivalente ao problema

$$\begin{array}{lll} (\tilde{D}) & \textit{minimizar} & \tilde{x}^T s \\ & \textit{sujeito a:} & P_A s = P_A c \\ & & s \geq 0, \end{array}$$

no sentido de que existe uma correspondência um-a-um entre $\mathcal{S}$ e $\tilde{\mathcal{S}}$ e entre $\mathcal{S}^0$ e $\tilde{\mathcal{S}}^0$, onde

$$\tilde{\mathcal{S}} = \{s; (y, s) \in \mathcal{S} \text{ para algum } y\}$$

e

$$\tilde{\mathcal{S}}^0 = \{s; (y, s) \in \mathcal{S}^0 \text{ para algum } y\},$$

preservando otimalidade. Assim, o problema (D) pode ser reescrito como um problema em s apenas. Veja página 8 em [56] e páginas 175 e 176 em [25].

No próximo capítulo enunciaremos e demonstraremos o Lema de Farkas, o qual será útil para a nossa estratégia de demonstrar o Teorema de Dualidade sem o auxílio de métodos de resolução de problemas de PL; em particular, o método simplex.

Atenção: Que tal fazer o exercício 9 agora?

7.3 Exercícios

1. Exibir dois subespaços vetoriais V e W em R^n que são ortogonais.

2. Seja a matriz 1×2 definida por $A = [1\ 1]$. Pede-se:

   (a) defina e desenhe $\mathcal{N}(A)$ e $\mathcal{R}(A^T)$; e

   (b) tome um vetor em $\mathcal{N}(A)$ e um outro em $\mathcal{R}(A^T)$ quaisquer, e verifique que estes vetores são ortogonais.

3. Seja a matriz 1×2 definida por $A = [1\ 1]$. Qual é o posto de A? Verifique que AA^T é não singular.

4. Considere a matriz $A = [1\ 1]$. Pede-se:

   (a) tome $c = [1, 2]^T$ e calcule c_p e $c_{\tilde{p}}$; e

(b) tome $c = [1, 1]^T$ e calcule c_p e $c_{\tilde{p}}$.

5. Existe alguma matriz de projeção que não é simétrica e nem idempotente? Justifique (você verificou a referência conforme Proposição 7.1.4?).

6. Considere o PPL

$$\begin{array}{llllllll}
\textit{minimizar} & & & 3x_5 & & & \\
\textit{sujeito a}: & & x_2 & +x_3 & -2x_4 & +x_5 & = 0 \\
& -x_1 & & & & -x_5 & \geq 0 \\
& -x_1 & & & +x_4 & -2x_5 & \geq 0 \\
& 2x_1 & & -x_3 & & +2x_5 & \geq 0 \\
& -x_1 & +x_2 & +2x_3 & -2x_4 & & = -3 \\
& & x_2, & x_3, & x_4 & \geq 0. &
\end{array}$$

Pede-se (para os itens (a) e (b) você pode consultar as regras de passagens do primal para o dual em vários livros de PL existentes nas bibliotecas!):

(a) forneça o seu dual;

(b) forneça o dual do dual; e

(c) qual a conclusão para o resultado do primeiro item (consulte a página 94 na referência [36])? E, qual a conclusão para o segundo item?

7. Seja o PPL no formato padrão definido por

$$\begin{array}{lll}
\textit{minimizar} & x_1 & \\
\textit{sujeito a}: & x_1 + x_2 + x_3 & = 3 \\
& x_1 + x_2 - x_3 & = 1 \\
& x_1, \quad x_2, \quad x_3 \geq 0. &
\end{array}$$

Pede-se:

(a) identificar A^T, b e c;

(b) definir o seu problema dual e desenhá-lo apenas para as variáveis duais; e

(c) para o PPL dual, identificar a função objetivo; os conjuntos viável, de pontos interiores, de pontos interiores viáveis e de soluções ótimas; alguns pontos viáveis, interiores, interiores viáveis uma solução ótima e o valor ótimo.

8. Para o exercício anterior, encontre uma folga dual viável $s \geq 0$ tal que $P_A s = P_A c$.

9. Para o exercício anterior, encontre uma solução para o problema dual apenas em s e, depois, encontre o valor único de $y \in R^2$.

Capítulo 8

O Lema de Farkas

Continuamos com o nosso estudo dos fundamentos da PL. Nosso objetivo aqui é enunciar e demonstrar o Lema de Farkas. Além disso, definimos e relacionamos direções viáveis com o espaço nulo da matriz tecnológica A. Faremos uso destes resultados no próximo capítulo acerca do Teorema de Dualidade.

Aqui, optamos em enunciar o Teorema de Separação, porém, sem uma demonstração. A razão desta nossa opção é não estendermos o nosso estudo para o campo da Análise Convexa.

8.1 O Lema

Iniciaremos o nosso propósito enunciando o Teorema de Separação, cuja demonstração pode ser encontrada nas páginas 45 e 46 em [6].

Teorema 8.1.1 *Seja C um conjunto convexo, fechado e não vazio em R^n e considere $z \notin C$. Então, existem um vetor $a \in R^n$, $a \neq 0$, e um escalar $\delta \in R$ tais que, para todo $x \in C$,*

$$a^T z > \delta \quad e \quad a^T x \leq \delta.$$

Deste teorema, dizemos que o hiperplano com a equação $a^T x = \delta$ separa z de C.

Atenção: Que tal fazer o exercício 1 agora?

Agora vamos definir um cone convexo gerado pelas colunas de uma matriz, que nos auxiliará no entendimento do Lema de Farkas.

Definição 8.1.2 *(a) Dizemos que um conjunto $C \subset R^n$ é um cone, quando para qualquer ponto $x \in C$ e para qualquer escalar não negativo λ, o ponto λx pertence a C.*

(b) Dada a matriz $A \in R^{m\times n}$, dizemos que o conjunto

$$\{w \in R^m; w = Az, z \geq 0\}$$

é o cone convexo gerado pelas colunas de A.

Atenção: Que tal fazer o exercício 2 agora?

O Teorema de Separação nos habilita a demonstrar o próximo resultado conhecido como o Lema de Farkas, cuja demonstração pode ser encontrada nas páginas III-36 e III-37 em [35].

Teorema 8.1.3 *Sejam A uma matriz $m \times n$ e um vetor b em R^m. Então, exatamente um dos dois seguintes sistemas tem uma solução:*

(1) $Ax = b$ e $x \geq 0$ para algum $x \in R^n$; e

(2) $A^T y \leq 0$ e $b^T y > 0$ para algum $y \in R^m$.

Demonstração: Suponhanhos que o sistema (1) tenha uma solução. Isto significa que existe um vetor $\hat{x} \geq 0$ tal que $A\hat{x} = b$. Fixe arbitrariamente $y \in R^m$ tal que $A^T y \leq 0$. Desta forma, segue-se que

$$b^T y = (A\hat{x})^T y = \hat{x}^T A^T y \leq 0.$$

Pela arbitrariedade de y, o sistema (2) não tem solução.

Agora, suponhamos que o sistema (1) não tenha solução. Defina o cone convexo gerado pelas colunas da matriz A por

$$S = \{y \in R^m; y = Ax,\ x \geq 0\}.$$

Por hipótese, $b \notin S$. Ainda, como S é um poliedro, segue-se pela Proposição 4.1.10 que S é um conjunto convexo e fechado. Também, S é não vazio, porque $0 \in S$. Pelo Teorema de Separação, existem um vetor não nulo $a \in R^m$ e um escalar $\delta \in R$ tais que, para todo $y \in S$, $a^T b > \delta$ e $a^T y \leq \delta$. Como $0 \in S$, $\delta \geq 0$ e, daí, $a^T b > 0$. Ainda, uma vez que $a^T y \leq \delta$, segue-se pela definição de S que, para todo $j = 1, \ldots, n$ e para qualquer número real $\lambda > 0$,

$$\delta \geq a^T \lambda A_j = \lambda e_j^T A^T a,$$

onde e_j é o j-ésimo vetor unitário em R^n. Dividindo esta última desigualdade por λ e tomando o limite quando $\lambda \to \infty$, segue-se que $A^T a \leq 0$. Portanto, como demonstramos que $A^T a \leq 0$ e $b^T a > 0$, para algum $a \in R^m$, concluímos que a é uma solução para o sitema (2). Isto finaliza a demonstração. ∎

Atenção: Que tal fazer o exercício 3 agora?

Na próxima seção estudaremos um pouco sobre direções viáveis.

8.2 Direções viáveis

Aqui, definimos e relacionamos direções viáveis com o espaço nulo da matriz $A \in R^{m \times n}$. Iniciamos o nosso propósito definindo direções viáveis.

Definição 8.2.1 *Seja C um subconjunto de R^n. Um vetor $u \in R^n$ é uma direção viável a partir de $x \in C$, quando existe $\bar{\lambda} > 0$, tal que para qualquer $\lambda \in [0, \bar{\lambda}]$, $x + \lambda u \in C$.*

Atenção: Que tal fazer o exercício 4 agora?

Agora, vamos relacionar esta última definição com vetores em $\mathcal{N}(A)$, onde A é a matriz associada ao PPL (P). Iniciamos com a seguinte proposição.

Proposição 8.2.2 *Considere o conjunto viável $\mathcal{X}$ de (P). Se $u \in R^n$ é uma direção viável a partir de $x \in \mathcal{X}$, então $u \in \mathcal{N}(A)$.*

Demonstração: Suponhamos que $u \in R^n$ é uma direção viável a partir de $x \in \mathcal{X}$. Por hipótese, existe $\bar{\lambda} > 0$ tal que para qualquer $\lambda \in [0, \bar{\lambda}]$, $x + \lambda u \in \mathcal{X}$. Segue-se, usando a linearidade de A, que

$$b = A(x + \lambda u) = Ax + \lambda Au \Rightarrow \lambda Au = 0,$$

onde a implicação decorre do fato de que $x \in \mathcal{X}$. Então, uma vez que podemos ter $\lambda > 0$, concluímos que $u \in \mathcal{N}(A)$; finalizando a demonstração. ∎

Atenção: Que tal fazer o exercício 5 agora?

A próxima proposição pode ser encontrada nas páginas 11 e 12 em [23], a qual relaciona direções viáveis com o espaço nulo de A.

Proposição 8.2.3 *Seja $\mathcal{X}^0 = \{x \in \mathcal{X}; x > 0\}$ o conjunto de pontos interiores viáveis. Então:*

(a) Se $x \in \mathcal{X}^0$, então $u \in R^n$ é uma direção viável a partir de x se, e somente se, $u \in \mathcal{N}(A)$.

(b) Se $x \in (\mathcal{X} - \mathcal{X}^0)$, então $u \in R^n$ é uma direção viável a partir de x se, e somente se, $u \in \mathcal{N}(A)$ e $u_j \geq 0$ para todo $j = 1, \ldots, n$ tal que $x_j = 0$.

Demonstração:

(a) Se $u \in R^n$ é uma direção viável a partir de $x \in \mathcal{X}$, então $u \in \mathcal{N}(A)$; conforme proposição anterior. Por outro lado, suponhamos $u \in \mathcal{N}(A)$. Então, usando o fato de que $x \in \mathcal{X}^0$, para qualquer $\lambda \in R$ tem-se que

$$A(x + \lambda u) = Ax + \lambda Au = b.$$

Assim, basta examinar a restrição de não negatividade. Se $u \geq 0$, então para qualquer $\lambda \geq 0$, $x + \lambda u \geq 0$. Se alguma componente de u é negativa, então definimos

$$\bar{\lambda}_i = \sup\{\lambda; x_i + \lambda u_i \geq 0,\ u_i < 0,\ i = 1, ..., n\}$$

e, para a não negatividade, se $u_i < 0$, então $x_i + \lambda u_i = 0$ e conseqüentemente $\bar{\lambda}_i = -x_i/u_i$. O maior passo que garante viabilidade é $\min\{\bar{\lambda}_i; u_i < 0,\ i = 1, \ldots, n\}$, o que finaliza a primeira parte da demonstração.

(b) Suponhamos que $u \in R^n$ é uma direção viável a partir de $x \in (\mathcal{X} - \mathcal{X}^0)$. Por hipótese, segue-se que $u_j \geq 0$ para todo $j = 1, \ldots, n$ tal que $x_j = 0$. Além disso, usando a proposição anterior, $u \in \mathcal{N}(A)$. Por outro lado, suponhamos que $u_j \geq 0$ para todo $j = 1, \ldots, n$ tal que $x_j = 0$ e

$u \in \mathcal{N}(A)$. Como $u \in \mathcal{N}(A)$, $A(x + \lambda u) = b$ para todo $\lambda \in R$. Resta mostrar que existe $\bar{\lambda} > 0$ tal que $x + \lambda u \geq 0$ para todo $\lambda \in [0, \bar{\lambda}]$. Para isto, definimos

$$J = \{1, \ldots, n\}, \ J_0 = \{j \in J; x_j = 0\} \text{ e } J_1 = J - J_0.$$

Considere $j \in J_0$ fixo, porém, arbitrário. Logo, $x_j + \lambda u_j \geq 0$, para todo $\lambda \geq 0$, uma vez que $u_j \geq 0$. Agora, considere $i \in J_1$ fixo, porém, arbitrário. Logo, se $u_i \geq 0$, então $x_i + \lambda u_i \geq 0$, para todo $\lambda \geq 0$. E, se $u_i < 0$, então usando o item (a), obtemos um tamanho de passo positivo. Pela arbitrariedade de $i \in J_1$, podemos usar o item (a) novamente obtendo

$$\bar{\lambda} = \min\{\frac{-x_i}{u_i}; u_i < 0, \ i = 1, ..., n\}.$$

Portanto, pela arbitrariedade de $j \in J_0$ e $i \in J_1$, $x + \lambda u \geq 0$ para todo $\lambda \in [0, \bar{\lambda}]$, como queríamos demonstrar. Isto finaliza a demonstração.

∎

No próximo capítulo estudaremos os teoremas de dualidade.

Atenção: Que tal fazer o exercício 6 agora?

8.3 Exercícios

1. Considere o conjunto $C = \{x \in R^2; \|x\| \leq 1\}$ e o ponto $z = [2, 2]^T$. Calcule dois hiperplanos que separam z de C.

2. Considere a matriz

$$A = \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \end{bmatrix}.$$

Dê um exemplo de um cone que não é convexo e dê um exemplo de um cone convexo gerado pelas colunas da matriz A.

3. Interprete graficamente o Lema de Farkas.

4. Seja o poliedro definido por

$$\{x \in R^2;\ x_1 + x_2 = 2,\ x_1,\ x_2 \geq 0\}.$$

   Pede-se:

   (a) encontre uma direção viável a partir de $x = [1, 1]^T$; e

   (b) encontre uma direção não viável a partir de $x = [1, 1]^T$.

5. Considere o item (a) do exercício anterior. Verifique que a sua direção viável está em $\mathcal{N}(A)$.

6. Seja o sistema $Ax = b$ definido por $A = [1\ \ -2]$ e $b = 2$ com $x \geq 0$. Pede-se:

   (a) encontre uma direção viável a partir de $x = [4, 1]^T$ e verifique que esta direção está em $\mathcal{N}(A)$; e

   (b) encontre uma direção viável $u \in R^2$ a partir de $x = [2, 0]^T$ e verifique que esta direção está em $\mathcal{N}(A)$ e que u_2 satisfaz a não negatividade.

Capítulo 9

Dualidade

Continuamos com o nosso estudo dos fundamentos da PL. Aqui estudamos o problema dual com o nosso objetivo centrado nos teoremas de dualidade. Além disso, concluímos que supor posto completo para a matriz A não representa perda de generalidade tanto para o problema primal (conforme Capítulo 2) quanto para o problema dual.

9.1 Três teoremas de dualidade

Nossa intenção agora é estabelecer o Teorema de Dualidade. Para isto vamos relacionar os problemas primal e dual entre si. Iniciamos com o Teorema de Dualidade Fraco, que relaciona os problemas primal e dual através de pontos viáveis, no sentido de que obtemos limitantes inferior e superior para os problemas (P) e (D), respectivamente. Sua demonstração pode ser encontrada na página 28 em [46].

Teorema 9.1.1 *Suponha $x \in \mathcal{X}$ e $(y, s) \in \mathcal{S}$ quaisquer. Então $c^T x \geq b^T y$.*

Demonstração: Por hipótese, segue-se diretamente que

$$c^T x - b^T y = (A^T y + s)^T x - (Ax)^T y = x^T s \geq 0;$$

finalizando a demonstração. ∎

Observe na demonstração acima que $c^T x - b^T y = x^T s$. Usualmente chamamos $c^T x - b^T y$ o *gap* de dualidade e $x^T s$ o *gap* de complementaridade.

Atenção: Que tal fazer o exercício 1 agora?

Agora vamos demonstrar o Teorema de Dualidade, que relaciona os problemas primal e dual entre si, no sentido de que ou um PPL possui uma solução ótima, ou é um problema ilimitado ou é um problema inviável. Sua demonstração pode ser encontrada nas páginas 71, 72 e 73 em [52].

Teorema 9.1.2 *Considere os problemas primal (P) e dual (D). Uma, e somente uma, das seguintes afirmações é correta:*

(a) se o problema (P) tem uma solução ótima, então o problema (D) também tem uma solução ótima e os valores das funções objetivos de ambos são iguais. Se o problema (D) tem uma solução ótima, então o problema (P) também tem uma solução ótima e os valores das funções objetivos de ambos são iguais.

(b) Se (P) é um problema ilimitado, então (D) é um problema inviável. Se (D) é um problema ilimitado, então (P) é um problema inviável.

Demonstração:

(a) Basta mostrar que se o problema (P) tem uma solução ótima, então o problema (D) também tem uma solução ótima e os valores das funções objetivos de ambos são iguais. O outro caso pode ser demonstrado de maneira análoga. Assim, suponha que x^* é uma solução ótima para o problema primal (P). Seja o sistema (1),

$$\begin{array}{rrrcl} Ax & & & = & b \\ & A^T y & +s & = & c \\ c^T x & -b^T y & & = & 0 \\ x, & & s & \geq & 0. \end{array}$$

Como

$$x^T s = x^T(c - A^T y) = c^T x - b^T y = 0$$

e, usando o Lema 7.2.2, este sistema é equivalente ao sistema (1'),

$$\begin{array}{rrcl} Ax & & = & b \\ & P_A s & = & P_A c \\ & x^T s & = & 0 \\ x, & s & \geq & 0. \end{array}$$

Reescrevendo este último sistema, considere $x^T s = 0$ tal que

$$\begin{array}{rcrcl} Ax & & & = & b \\ & & -P_A s & = & -P_A c \\ x, & & s & \geq & 0. \end{array}$$

Defina

$$\bar{A} = \begin{bmatrix} A & 0 \\ 0 & -P_A \end{bmatrix}, \ \bar{b} = \begin{bmatrix} b \\ -P_A c \end{bmatrix} \text{ e } w = \begin{bmatrix} x \\ s \end{bmatrix} \in R^{2n}.$$

Se o sistema $\bar{A}w = \bar{b}$ e $w \geq 0$, com $x^T s = 0$, tem solução, então pela hipótese de x^* e usando o Teorema de Dualidade Fraco, este teorema é satisfeito. Senão, podemos usar o Lema de Farkas, obtendo o sistema (2) abaixo que admite solução,

$$\bar{A}^T z \leq 0 \text{ e } \bar{b}^T z > 0.$$

Nossa estratégia agora é demonstrar que o sistema (2) não pode admitir solução. Para algum $z^T = [u^T \ v^T] \in R^{m+n}$, podemos reescrever o sistema (2) como o sistema (2'), a saber:

$$\begin{array}{rrcl} A^T u & & \leq & 0 \\ & -P_A^T v & \leq & 0 \\ b^T u & -(P_A c)^T v & > & 0. \end{array}$$

Se $(P_A c)^T v \geq 0$, então $b^T u > 0$, onde por hipótese, $b = Ax^*$ e $x^* \geq 0$. Além disso, $A^T u \leq 0$ por (2'). Logo,

$$0 < b^T u = (Ax^*)^T u = (x^*)^T A^T u \leq 0.$$

Ou seja, $(P_A c)^T v \geq 0$ não pode ocorrer. Finalmente, se $(P_A c)^T v < 0$, vejamos:

$$(P_A c)^T v = c^T P_A^T v = c^T (P_A v) = c^T v_p,$$

onde as segunda e terceira igualdades decorrem, respectivamente, da Proposição 7.1.5 e da Proposição 7.1.4. Então, v_p é uma direção viável, a partir de algum ponto viável $\hat{x}$ do problema primal, conforme Proposição 8.2.3, uma vez que $P_A v = v_p \geq 0$ por (2') e $v_p \in \mathcal{N}(A)$. Logo, o vetor $x(\lambda) = \hat{x} + \lambda v_p$ é um ponto viável do problema primal com $c^T x(\lambda) = c^T \hat{x} + \lambda c^T v_p \to -\infty$, quando $\lambda \to \infty$. Isto contradiz a hipótese de x^*. Portanto, o sistema (2) não admite solução. Novamente, pelo Lema de Farkas, o sistema (1) de fato tem solução. Isto finaliza a demonstração do item (a).

(b) Basta mostrar que se (P) é um problema ilimitado, então (D) é um problema inviável. O outro caso pode ser demonstrado de maneira análoga. Assim, supomos que (P) é um problema ilimitado. Logo, por definição, existe uma seqüência (x^k) tal que $x^k \in \mathcal{X}$ e $c^T x^k \to -\infty$. Desta forma, $\mathcal{X} \neq \emptyset$. Suponha por contradição que $\mathcal{S} \neq \emptyset$. Então, pelo Teorema de Dualidade Fraco, para quaisquer $x \in \mathcal{X}$ e $(y, s) \in \mathcal{S}$, temos $c^T x \geq b^T y$. Isto contradiz o fato de que (P) é um problema ilimitado; concluindo a demonstração do item (b).

∎

Observe que se (P) (ou (D)) é inviável então o seu problema dual pode ser ilimitado ou inviável.

Atenção: Que tal fazer o exercício 2 agora?

Agora pretendemos relacionar os problemas primal (P) e dual (D) através de soluções ótimas. O resultado que nos possibilita esta associação é o Teorema de Dualidade Forte, cuja demonstração pode ser encontrada na página 193 em [44].

Teorema 9.1.3 *Suponha que os problemas (P) e (D) têm soluções viáveis. Então, ambos têm soluções ótimas $x^* \in \mathcal{X}(P)$ e $(y^*, s^*) \in \mathcal{S}(D)$ e, necessariamente, $c^T x^* - b^T y^* = (x^*)^T s^* = 0$.*

Demonstração: Por hipótese e usando o Teorema de Dualidade Fraco, para todo $x \in \mathcal{X}$ e para todo $(y, s) \in \mathcal{S}$, $c^T x \geq b^T y$. Então, pelo Teorema de Dualidade, existe $x^* \in \mathcal{X}(P)$. Finalmente, pelo item (a) do Teorema de Dualidade, o resultado se segue, finalizando a demonstração. ∎

Atenção: Que tal fazer o exercício 3 agora?

Finalizamos este capítulo concluindo que a hipótese de que a matriz A tem posto completo é, de fato, sem perda de generalidade.

9.2 Sobre a hipótese do posto de A

Considere a transformação linear definida pela matriz $A \in R^{m \times n}$. No Capítulo 7 definimos dois subespaços importantes do espaço vetorial R^n associados com esta transformação: o espaço nulo de A e o espaço linha de A. Agora, definiremos dois subespaços importantes do espaço vetorial R^m associados com esta transformação: o espaço nulo à esquerda de A, definido por

$$\mathcal{N}(A^T) = \{y \in R^m; A^T y = 0\}$$

e seu complemento ortogonal, o espaço coluna de A, definido por

$$\mathcal{R}(A) = \{y \in R^m; y = Aw, w \in R^n\}.$$

Atenção: Que tal fazer o exercício 4 agora?

Consideremos o problema dual (D). Se as colunas da matriz A^T são linearmente dependentes, isto é, existe um vetor não nulo $y \in \mathcal{N}(A^T)$ tal que $A^T y = 0$ e, além disso, $b^T y \neq 0$, isto é, $b \notin \mathcal{R}(A)$, o que significa que o problema primal é um problema inviável, então, pelo Teorema de Dualidade, se $\mathcal{S}$ é não vazio, o problema (D) é ilimitado. Se $b^T y = 0$ para todo y, então podemos eliminar uma das colunas correspondentes a uma componente não nula de y sem afetar (D). Continuando com este processo obtemos um problema equivalente onde A tem posto completo. Desta forma, supor que A tem posto completo não perde generalidade neste caso, também; conforme vimos para o problema primal (P) no Capítulo 2. Veja página 8 em [56].

No próximo capítulo estudaremos o par de problemas primal e dual, isto é, o problema primal-dual.

Atenção: Que tal fazer o exercício 5 agora?

9.3 Exercícios

1. Suponha $x^* \in \mathcal{X}$ e $(y^*, s^*) \in \mathcal{S}$ quaisquer. Demonstre que, se

$$c^T x^* - b^T y^* = (x^*)^T s^* = 0,$$

então $x^* \in \mathcal{X}(P)$ e $(y^*, s^*) \in \mathcal{S}(D)$.

2. Dê exemplos para:
   (a) um problema primal inviável e o seu dual ilimitado; e
   (b) um problema primal inviável e o seu dual inviável.

3. Seja o PPL no formato padrão definido por

$$A = [1\ 1\ 1],\ b = 3 \text{ e } c = [1, 0, 0]^T.$$

   Considere um ponto viável primal dado por $x = [1, 1, 1]^T$ e um ponto viável dual dado por $(y, s^T) = [-1, 2, 1, 1]$. Encontre uma solução ótima primal, uma solução ótima dual e o valor da função objetivo para ambos os problemas.

4. Seja uma matriz definida por $A = [1\ 1]$. Pede-se:
   (a) defina $\mathcal{N}(A^T)$ e $\mathcal{R}(A)$; e
   (b) tome um vetor em $\mathcal{N}(A^T)$ e um outro em $\mathcal{R}(A)$ quaisquer, e verifique que estes vetores são ortogonais.

5. Seja o PPL no formato padrão definido por

$$\begin{array}{ll} \textit{minimizar} & x_1 \\ \textit{sujeito a}: & x_1 + x_2 + x_3 \qquad = 2 \\ & x_1 + x_2 + x_3 \qquad = 2 \\ & x_1, \quad x_2, \quad x_3 \geq 0. \end{array}$$

   Pergunta-se:
   (a) a matriz A^T tem colunas linearmente dependentes? Justifique; e
   (b) definindo o problema dual, podemos eliminar a segunda coluna de A^T? Justifique.

Capítulo 10

O Problema Primal-Dual

Finalizamos o nosso estudo dos fundamentos em PL estudando o problema primal-dual, com o nosso objetivo centrado nas condições de otimalidade. Além disso, demonstramos o Teorema de Complementaridade Estrita e estudamos aspectos relevantes sobre a geometria do par de problemas primal e dual.

10.1 O problema

Consideremos os problemas primal (P) e dual (D).

Iniciamos o nosso estudo para definir o problema primal-dual com a seguinte definição.

Definição 10.1.1 *Sejam $x^* \in \mathcal{X}$ e $s^* \in \tilde{\mathcal{S}}$ soluções ótimas para os problemas primal e dual, respectivamente.*

(a) A igualdade $(x^*)^T s^* = 0$ é chamada condição de folga complementar.

(b) Dizemos que (x^*, s^*) é um par de soluções complementares, quando (x^*, s^*) satisfaz a condição de folga complementar.

Atenção: Que tal fazer o exercício 1 agora?

A seguir, demonstraremos uma proposição bastante útil para o nosso estudo para definir o problema primal-dual. Sua demonstração pode ser encontrada na página 200 em [26].

Proposição 10.1.2 *Considere os vetores* $x, s \in R^n_+$. *Temos que,* $xs = 0$ *se, e somente se,* $x^T s = 0$.

Demonstração: Temos que, $xs = 0$ significa que $x_j s_j = 0$, $j = 1, \ldots, n$. Isto é equivalente a $x^T s = \sum_{j=1}^n x_j s_j = 0$, para $(x, s) \geq 0$. Isto finaliza a demonstração. ∎

Atenção: Que tal fazer o exercício 2 agora?

O próximo teorema é também conhecido como o Teorema das Folgas Complementares, cuja demonstração pode ser encontrada na página 75 em [52].

Teorema 10.1.3 *Considere* $x^* \in \mathcal{X}$ *e* $s^* \in \tilde{\mathcal{S}}$ *soluções ótimas, respectivamente, para os problemas primal e dual. Então* $x^* s^* = 0$.

Demonstração: Pelo Teorema de Dualidade $(x^*)^T s^* = 0$. Pela proposição anterior, o resultado se segue, finalizando a demonstração. ∎

Considere $x^* \in \mathcal{X}$ e $s^* \in \tilde{\mathcal{S}}$ soluções ótimas, respectivamente, para os problemas primal e dual. Este teorema afirma que a igualdade $x^* s^* = 0$ tem estrutura combinatorial: para cada $j = 1, \ldots, n$, $x_j^* = 0$ ou $s_j^* = 0$.

Atenção: Que tal fazer o exercício 3 agora?

O problema primal-dual é definido assim: dados uma matriz A, $m \times n$, $0 < m < n$, $posto(A) = m$, e vetores $b \in R^m$ e $c \in R^n$, encontrar, se existir, uma solução para o sistema de equações e inequações

$$(PD) \quad \begin{array}{rcl} Ax & = & b \\ A^T y + s & = & c \\ xs & = & 0 \\ x, s & \geq & 0. \end{array}$$

Usando a Proposição 10.1.2, o sistema não linear (PD) pode ser visto como um sistema linear, bastando substituir $xs = 0$ por

$$x^T s = c^T x - b^T y = 0.$$

Observe que o problema para encontrar $x \in R^n$ que cumpre as condições $Ax = b$ e $x \geq 0$, é o problema de viabilidade primal. Ainda, o problema para encontrar $(y, s) \in R^m \times R^n$ que cumpre as condições $A^T y + s = c$ e $s \geq 0$, é o problema de viabilidade dual. Além disso, usando a Proposição 10.1.2 e o Teorema de Dualidade, encontrar $(x, s) \in \mathcal{X} \times \tilde{\mathcal{S}}$ que cumpre a condição $xs = 0$, é encontrar um par de soluções complementares.

Definimos o conjunto viável primal-dual por

$$\mathcal{F} = \{(x, s) \in R^n \times R^n; x \in \mathcal{X}, s \in \tilde{\mathcal{S}}\},$$

o conjunto viável de pontos interiores primal-dual por

$$\mathcal{F}^0 = \{(x, s) \in \mathcal{F}; (x, s) > 0\}$$

e o conjunto de soluções ótimas primal-dual por

$$\mathcal{F}(PD) = \{(x, s) \in \mathcal{F}; xs = 0\}.$$

Atenção: Que tal fazer o exercício 4 agora?

As condições de otimalidade para o par de problemas primal e dual estão expressas no próximo teorema.

Teorema 10.1.4 *Considere os problemas primal* (P) *e dual* (D)*. Um ponto* $x \in R^n$ *é uma solução ótima de* (P) *se, e somente se, existe um par de multiplicadores (de Lagrange)* $(y, s) \in R^m \times R^n$*, tal que o sistema*

$$\begin{array}{rcl} Ax & = & b \\ A^T y + s & = & c \\ x^T s & = & 0 \\ x, s & \geq & 0 \end{array}$$

é satisfeito.

Demonstração: Imediata usando o Teorema de Dualidade. ∎

As condições de otimalidade para o par de problemas primal e dual coincidem com as condições de Karush-Kuhn-Tucker, a saber: x é uma solução ótima de (P) se, e somente se, existe um par de multiplicadores de Karush-Kuhn-Tucker (y, s) tal que o sistema (PD) é satisfeito.

Este resultado é um caso particular do Teorema de Karush-Kuhn-Tucker para Programação Não Linear.

Além disso, pelo teorema anterior, as condições de otimalidade para um PPL consiste em encontrar um ponto viável x em $\mathcal{X}$ e um outro s em $\tilde{\mathcal{S}}$ com, necessariamente, $x^T s = 0$. Como o *gap* de complementaridade é sempre não negativo e sempre igual ao *gap* de dualidade $c^T x - b^T y$, podemos definir o problema (PD) como um PPL, a saber:

$$\begin{array}{ll} \textit{minimizar} & c^T x - b^T y \\ \textit{sujeito a:} & Ax = b \\ & A^T y + s = c \\ & x, \quad s \geq 0. \end{array}$$

Atenção: Que tal fazer o exercício 5 agora?

Na próxima seção enunciaremos e demonstraremos o Teorema das Folgas Complementares Estritas, ou simplesmente Teorema de Complementaridade Estrita.

10.2 O Teorema de Complementaridade Estrita

O Teorema de Complementaridade Estrita é também conhecido como o Teorema de Goldmann e Tucker. Para estudá-lo, iniciamos com a seguinte definição.

Definição 10.2.1 *Sejam $x^* \in \mathcal{X}$ e $s^* \in \tilde{\mathcal{S}}$ soluções ótimas para os problemas primal e dual, respectivamente.*

(a) A desigualdade estrita $x^* + s^* > 0$ é chamada condição de folga complementar estrita.

(b) Dizemos que (x^*, s^*) é um par de soluções complementares estritas, quando (x^*, s^*) satisfaz a condição de folga complementar estrita.

Atenção: Que tal fazer o exercício 6 agora?

Agora estamos prontos para enunciar e demonstrar o Teorema de Complementaridade Estrita, que afirma que qualquer PPL com uma solução ótima possui um par de soluções complementares estritas. Sua demonstração pode ser encontrada nas páginas 77 e 78 em [52].

Teorema 10.2.2 *Suponha $\mathcal{X}$ e $\tilde{\mathcal{S}}$ conjuntos não vazios. Então, os problemas primal e dual têm um par de soluções complementares estritas x^* e s^*, isto é, $x^* + s^* > 0$.*

Demonstração: Por hipótese, podemos usar o Teorema de Dualidade Forte. Denotamos $\gamma = v(P) = v(D)$ o valor ótimo. Inicialmente, queremos demonstrar que $x_j^* = 0$ para toda solução ótima para o problema primal se, e somente se, $s_j^* > 0$ para alguma solução ótima para o problema dual. Se $s_j^* > 0$ para alguma solução ótima para o dual, então pelo Teorema 10.1.3, $x_j^* = 0$ para toda solução ótima para o primal. Assim, suponha $x_j^* = 0$ para toda solução ótima para o primal. Vamos demonstrar que $s_j^* > 0$ para alguma solução ótima para o dual. Com efeito, $x_j^* = 0$ é equivalente ao problema (1)

$$\begin{array}{lll} \textit{minimizar} & -(u^j)^T x & \\ \textit{sujeito a:} & \quad Ax & = b \\ & -c^T x - t & = -\gamma \\ & x \geq 0,\ t \geq 0 & \end{array}$$

admitir uma solução ótima $(\hat{x}, \hat{t})$ com valor ótimo $-(u^j)^T \hat{x} = 0$, onde $u^j \in R^n$ é um vetor de zeros com a j-ésima coordenada igual a 1. Equivalentemente, usando o Teorema de Dualidade, o problema dual (2) de (1),

$$\begin{array}{lll} \textit{maximizar} & b^T y - \gamma\lambda & \\ \textit{sujeito a:} & A^T y - c\lambda + s & = -u^j \\ & \quad -\lambda + s_{n+1} & = 0 \\ & s \geq 0,\ s_{n+1} \geq 0 & \end{array}$$

também admite solução ótima $(\hat{y}, \hat{\lambda}, \hat{s}, \hat{s}_{n+1})$ com valor ótimo

$$b^T \hat{y} - \gamma\hat{\lambda} = -(u^j)^T \hat{x} = 0.$$

Defina $\hat{\lambda} = \hat{s}_{n+1} \geq 0$. Agora, considere $(\bar{y}, \bar{s})$ uma solução ótima para o problema dual (D). Segue-se que

$$(A^T\bar{y} + \bar{s}) + (A^T\hat{y} - c\hat{\lambda} + \hat{s}) = c - u^j.$$

Daí,

$$\bar{s} + \hat{s} + u^j = (1 + \hat{\lambda})c - A^T(\bar{y} + \hat{y}).$$

Dividindo ambos os lados por $1 + \hat{\lambda} > 0$,

$$\frac{\bar{s} + \hat{s} + u^j}{1 + \hat{\lambda}} = c - A^T \frac{(\bar{y} + \hat{y})}{1 + \hat{\lambda}}.$$

Tomando

$$y^* = \frac{(\bar{y}+\hat{y})}{1+\hat{\lambda}} \quad \text{e} \quad s^* = \frac{\bar{s}+\hat{s}+u^j}{1+\hat{\lambda}} \geq 0,$$

segue-se que

$$c - A^T y^* = s^* \geq 0 \;\; \text{e} \;\; s_j^* = \frac{\bar{s}_j + \hat{s}_j + 1}{1+\hat{\lambda}} \geq \frac{1}{1+\hat{\lambda}} > 0.$$

Então, para todo $\tilde{x} \in \mathcal{X}(P)$,

$$\tilde{x}^T s^* = \tilde{x}^T(c - A^T y^*) = \gamma - b^T y^* = \gamma - b^T \frac{\bar{y}+\hat{y}}{1+\hat{\lambda}} = 0.$$

Logo, concluímos que se $x_j^* = 0$ para toda solução ótima para o primal, então $s_j^* > 0$ para alguma solução ótima para o dual. Finalmente, devemos exibir $x_i^* > 0$ para $s_i^* = 0$. Pelo Teorema 6.2.3 $\mathcal{X}(P)$ é um poliedro, logo, convexo. Então podemos definir

$$J = \{i = 1, \ldots, n; \exists x^i \in \mathcal{X}(P),\; x_i^i > 0\},$$

e tomar

$$x_i^* = \frac{1}{|J|} \sum_{i \in J} x_i^i > 0.$$

Isto finaliza a demonstração. ∎

Atenção: Que tal fazer o exercício 7 agora?

Uma conseqüência importante deste teorema é o próximo resultado, cuja demonstração pode ser encontrada na página 9 em [56].

Proposição 10.2.3 *Se c não está no espaço linha de A, então $c^T x > v(P)$ para todo $x \in \mathcal{X}^0$.*

Demonstração: Suponha por contradição que $c^T x = v(P)$ para algum $x \in \mathcal{X}^0$. Então, (P) tem uma solução ótima e, pelo Teorema de Dualidade, (D) também tem uma solução ótima. Usando o Teorema das Folgas Complementares isto implica que $s^* = 0$ para qualquer solução ótima (y^*, s^*) em (D). Portanto, $c = A^T y^*$. Isto contradiz o fato de que c não está no espaço linha de A, finalizando a demonstração. ∎

Atenção: Que tal fazer os exercícios 8 e 9 agora?

Na próxima seção estudaremos alguns aspectos da geometria do par de problemas primal e dual.

10.3 Geometria

O Teorema Fundamental da PL afirma que uma solução ótima, se existir, é atingida em ao menos um ponto extremo. O que pretendemos estudar agora é a possibilidade de uma solução ótima, para o problema (PD), ser atingida em um par de soluções complementares estritas, no sentido de que estas soluções geram uma solução interior em um subconjunto do conjunto de soluções ótimas do problema (PD).

Iniciamos o nosso propósito com um resultado que nos auxiliará na tarefa de formular hipóteses para os problemas primal e dual, cuja demonstração pode ser encontrada nas páginas 9 e 10 em [56].

Teorema 10.3.1 *Suponha que $\mathcal{X}$ é não vazio. Então as seguintes afirmações são equivalentes:*

(a) $\mathcal{S}^0$ é não vazio;

(b) Para todo $\kappa \in R$, $\{x \in \mathcal{X}; c^T x \leq \kappa\}$ é limitado;

(c) $\mathcal{X}(P)$ é não vazio e limitado; e

(d) Para algum $\kappa \in R$, $\{x \in \mathcal{X}; c^T x \leq \kappa\}$ é não vazio e limitado.

Demonstração: $(a) \Rightarrow (b)$: Suponha $(y, s) \in \mathcal{S}^0$. Então, usando dualidade, $x \in \mathcal{X}$ e $c^T x \leq \kappa$ implica que

$$x^T s = c^T x - b^T y \leq \kappa - b^T y.$$

Daí, tais pontos x devem pertencer ao conjunto limitado

$$\{x; x^T s \leq \kappa - b^T y, x \geq 0\} \subseteq \{x; 0 \leq x \leq (\kappa - b^T y)s^{-1}\}.$$

$(b) \Rightarrow (c)$: Tomando $\kappa = c^T \bar{x}$ para algum $\bar{x} \in \mathcal{X}$, temos que (P) é equivalente a minimizar $c^T x$ sobre o conjunto compacto e não vazio

$$\{x \in \mathcal{X}; c^T x \leq \kappa\},$$

tal que ele possui uma solução ótima. Além disso, toda solução ótima pertence a este conjunto compacto.

$(c) \Rightarrow (d)$: Isto é imediato tomando $\kappa = v(P)$.

$(d) \Rightarrow (a)$: Suponha que $\{x \in \mathcal{X}; c^T x \leq \kappa\}$ é não vazio e limitado. Escolha algum $\bar{s} > 0$. Então o problema

$$\min\{c^T x; Ax = 0, x^T \bar{s} = 1, x \geq 0\}$$

é inviável ou tem valor ótimo positivo. Assim, seu dual

$$\max\{\zeta; A^T y + \bar{s}^T \zeta + s = c, s \geq 0\},$$

que claramente tem uma solução viável (tome $y = 0$ e ζ suficientemente negativo), tem uma solução viável com ζ positivo. Isto fornece um ponto em $\mathcal{S}^0$. ∎

Atenção: Que tal fazer o exercício 10 agora?

Agora vamos formalizar, no próximo teorema, o resultado que afirma que todo PPL possui uma solução ótima que satisfaz a condição de complementaridade estrita e gera uma solução interior em um subconjunto do conjunto de soluções ótimas do problema (PD). Sua demonstração pode ser encontrada na página 78 em [52].

Teorema 10.3.2 *Considere um PPL definido por (PD). Uma solução ótima para o problema (PD) satisfazendo a condição de folga complementar estrita, é um ponto interior em um subconjunto do conjunto de soluções ótimas de (PD).*

Demonstração: Considere $(x^*, s^*) \in \mathcal{F}(PD)$ satisfazendo a condição de folga complementar estrita. Definimos

$$J_1 = \{j = 1, \ldots, n; x_j^* = 0\} \text{ e } J_2 = \{j = 1, \ldots, n; s_j^* = 0\}.$$

Pelo Teorema 10.1.3, $(x, s) \in \mathcal{F}(PD)$ se, e somente se, $(x, s) \in \mathcal{F}$ tal que $x_j = 0$ para todo $j \notin J_2$ e $s_j = 0$ para todo $j \notin J_1$. Isto significa que (x, s) resolve o sistema

$$\begin{aligned} A_{J_2} x_{J_2} &= b \\ P_{A_{J_1}} s_{J_1} &= P_{A_{J_1}} c_{J_1} \\ x_{J_2} \geq 0, s_{J_1} \geq 0. & \end{aligned}$$

As soluções de complementaridade estrita (x^*, s^*) geram um ponto interior para este sistema. Isto finaliza a demonstração. ∎

A partir do próximo capítulo, estaremos interessados no estudo de algoritmos para resolver problemas de PL. Iniciaremos esta terceira etapa com o estudo do método simplex. Assim, neste momento, estamos finalizando a etapa sobre os fundamentos da PL.

Atenção: Que tal fazer o exercício 11 agora?

10.4 Exercícios

1. Dê um exemplo de um PPL que admite solução ótima, inclusive para o seu dual. Calcule um par de soluções complementares.

2. Considere os vetores $x = [1, 0]^T$ e $s = [0, 1]^T$. Calcule $x^T s$ e xs.

3. Encontre um par de soluções complementares para o PPL

$$\begin{array}{lll} \textit{minimizar} & -x_1 - x_2 & \\ \textit{sujeito a:} & x_1 + x_2 + x_3 & = 1 \\ & x_1, \quad x_2, \quad x_3 \geq 0. & \end{array}$$

4. Desenhe no R^2 o conjunto viável $\mathcal{X}$ e o conjunto viável $\tilde{\mathcal{S}}$ para o PPL primal-dual

$$\begin{array}{llll} x_1 - x_2 & & & = 2 \\ & y & +s_1 & = 1 \\ & -y & +s_2 & = 0 \\ & x_1 s_1 & & = 0 \\ & x_2 s_2 & & = 0 \\ x_1, & x_2, & s_1, s_2 & \geq 0. \end{array}$$

   Encontre sua solução ótima.

5. Demonstre que o conjunto viável $\mathcal{X}$ é limitado se, e somente se, o conjunto viável $\tilde{\mathcal{S}}$ é ilimitado. E, vice-versa.

6. Dê um exemplo de um PPL que admite solução ótima, inclusive para o seu dual. Calcule um par de soluções complementares estritas.

7. Faça uma pesquisa bibliográfica e observe que o Teorema de Complementaridade Estrita não vale para o problema de Programação Não Linear $\min\{x^2; x \geq 0\}$.

8. Desenhe no R^2 um exemplo para a Proposição 10.2.3.

9. Considere o problema primal (P) com $\mathcal{X}(P)$ não vazio. Demonstre que $\mathcal{X} = \mathcal{X}(P)$ se, e somente se, $c \in \mathcal{R}(A^T)$.

10. Dê um exemplo de um PPL com $\mathcal{X}$ e $\mathcal{S}^0$ não vazios e verifique os itens (b), (c) e (d) do Teorema 10.3.1.

11. Considere o PPL do exercício 3. Encontre um par de soluções complementares estritas.

Capítulo 11

Método Simplex: algoritmo mestre

Aqui e nos próximos quatro capítulos, trataremos de duas famílias de métodos para resolver problemas de PL: simplex e afim-escala.

Todo método advém da necessidade de resolvermos algum problema. Neste e nos próximos dois capítulos, estaremos interessados na solução dos problemas de PL através do estudo do método simplex, devido a Dantzig [13]. O objetivo destes três capítulos é enunciar e demonstrar que o algoritmo simplex de duas fases com a regra de Bland [9] converge.

Iniciamos o nosso propósito introduzindo um algoritmo mestre.

11.1 Algoritmo mestre

Consideremos o PPL primal no formato padrão

$$
\begin{array}{lll}
(P) & \textit{minimizar} & c^T x \\
 & \textit{sujeito a:} & Ax = b \\
 & & x \geq 0,
\end{array}
$$

onde são dados uma matriz $A \in R^{m \times n}$ e vetores $b \in R^m$ e $c \in R^n$, com $0 < m < n$.

Sem perda de generalidade, consideremos a matriz A de posto completo e o vetor do lado direito $b \geq 0$. Neste último caso, para alguma coordenada do vetor b negativa, basta multiplicar a equação correspondente por menos um. No primeiro caso, agimos conforme Capítulo 2.

Considere o problema (P). A idéia do método simplex baseia-se no Corolário 5.1.3, afirmando que o conjunto viável de (P) tem um número finito de pontos extremos; no Corolário 5.2.2, afirmando que o conjunto viável de (P) é um poliedro com ao menos um ponto extremo, quando não vazio; no Teorema Fundamental da PL (Teorema 6.1.1), afirmando que se (P) admite solução ótima, então uma solução ótima é atingida em ao menos um ponto extremo do conjunto viável de (P); e, no Teorema 5.1.2, que caracteriza ponto extremo através de solução básica viável.

Quer dizer, a idéia 'ingênua' do método simplex consiste em caminhar pela fronteira de um conjunto poliedral de um PPL (P), através de pontos extremos adjacentes sucessivos com valores da função objetivo estritamente decrescentes.

A seguir, enunciamos um algoritmo mestre em uma tentativa de exprimir algoritmicamente as idéias do método simplex.

Algoritmo 11.1.1 *Mestre.*

Dados: x^0 *solução básica viável inicial associada a uma matriz base inicial* B_0.

$k := 0$.

REPITA

Escolha, se possível, uma nova variável básica daquelas variáveis não básicas.

Escolha, se possível, uma nova variável não básica daquelas variáveis básicas.

Atualize B_{k+1} *e* x^{k+1}.

$k := k + 1$.

ATÉ QUE 'convirja'.

Atenção: Que tal fazer o exercício 1 agora?

No restante desta seção e, nos próximos dois capítulos, construiremos o algoritmo simplex baseado neste algoritmo mestre; finalizando com um teorema de convergência.

Nossa primeira pergunta é a seguinte:

como determinar uma solução básica viável inicial?

Denominamos este problema de fase 1 ou, equivalentemente, problema de viabilidade.

11.1.1 Fase 1

Consideremos o problema de PL (P). Para o método simplex, o problema fase 1 tem a seguinte forma:

$$\begin{array}{lll} (P_1) & \textit{minimizar} & e^T x_a \\ & \textit{sujeito a:} & Ax + x_a = b \\ & & x \geq 0, \quad x_a \geq 0, \end{array}$$

onde $x_a \in R^m$ é um vetor de variáveis chamadas variáveis artificiais e e é um vetor de uns em R^m.

Uma vez que $b \geq 0$, segue-se que, para o problema (P_1),

$$\begin{bmatrix} x \\ x_a \end{bmatrix} = \begin{bmatrix} 0 \\ b \end{bmatrix} \in R^{n+m}$$

é uma solução básica viável associada a uma matriz base, dada pela matriz identidade I, $m \times m$.

O resultado a seguir garante a existência de uma solução ótima para o problema fase 1, no formato (P_1). Ou seja, este problema de viabilidade jamais será um problema ilimitado.

Teorema 11.1.2 *O problema (P_1) admite solução ótima.*

Demonstração: Sabemos que o conjunto viável de (P_1), denotado nesta demonstração por $\mathcal{X}_1$,

$$\mathcal{X}_1 = \{z = [x^T, x_a^T]^T;\ Ax + x_a = b,\ z \geq 0\},$$

é não vazio, pois $[0^T, b^T]^T \in \mathcal{X}_1$. O conjunto

$$\{z \in \mathcal{X}_1;\ 0 \leq e^T x_a \leq e^T b\}$$

é limitado e não vazio. Segue-se pelo Teorema 10.3.1 que o conjunto de soluções ótimas de (P_1) é não vazio, finalizando a demonstração. ∎

Atenção: Que tal fazer o exercício 2 agora?

O próximo resultado é um critério de inviabilidade para o PPL original (P).

Teorema 11.1.3 *Considere $[\hat{x}^T, \hat{x}_a^T]^T$ uma solução ótima de (P_1). Então, (P) é um problema inviável se, e somente se, $\hat{x}_a \neq 0$.*

Demonstração: Suponha $[\hat{x}^T, \hat{x}_a^T]^T$ uma solução ótima de (P_1). Vamos demonstrar inicialmente que se (P) é um problema inviável, então $\hat{x}_a \neq 0$. Suponha por contradição que $\hat{x}_a = 0$. Então $\hat{x} \geq 0$ e $A\hat{x} = b$, o que significa que $\mathcal{X}$ é não vazio. Ou seja, (P) não é um problema inviável, o que é uma contradição. Finalmente, vamos demonstrar que se $\hat{x}_a \neq 0$, então (P) é inviável. Suponha por contradição que (P) não é um problema inviável. Isto significa que existe $[x^T, x_a^T]^T \in R_+^{n+m}$ tal que $Ax = b$ e $x_a = 0$. Assim, temos que $e^T\hat{x}_a \leq e^T x_a = 0$, o que contradiz o fato de que $\hat{x}_a \neq 0$. Isto completa a demonstração. ∎

Atenção: Que tal fazer o exercício 3 agora?

Uma resposta auxiliar para a nossa primeira pergunta de como determinar uma solução básica viável inicial para o problema original (P) é: resolver o problema fase 1. O procedimento para resolver o problema (P_1) é análogo àquele para a próxima fase, isto é, o problema de otimalidade, denominado fase 2.

Observe que o problema fase 1 foi construído de tal maneira que uma solução básica viável inicial está disponível e, pelo Teorema 11.1.2, possui uma solução ótima. Pelo Teorema 11.1.3, de duas uma, ou certificamos que (P) é um problema inviável ou uma solução básica viável inicial para a próxima fase pode ser obtida.

Como certificar que (P) é um problema inviável? Pelo Teorema 11.1.3, quando o vetor de variáveis artificiais $\hat{x}_a$, em qualquer solução ótima de (P_1), é não nulo. Em outras palavras, quando o valor ótimo para o problema (P_1) é estritamente positivo.

A resposta definitiva para a nossa primeira pergunta será desenvolvida na próxima subseção.

11.1.2 Transição: da fase 1 para a fase 2

Supondo que o problema original (P) não é um problema inviável, devemos fornecer uma solução básica viável inicial para o problema fase 2, isto é, para o problema original (P). Nesta subseção respondemos como fazê-lo.

Se todas as variáveis artificiais são variáveis não básicas na solução ótima para o problema (P_1), então basta eliminar as variáveis artificiais. Neste caso, a partir da fase 1 concluída, determinamos uma solução básica viável inicial para a próxima fase. Este é o caso mais simples.

O caso mais delicado é quando algumas das variáveis artificiais, nulas na solução ótima para (P_1), são variáveis básicas. Inicialmente, desenvolveremos este caso através de um exemplo.

Exemplo 11.1.4 *Considere o PPL*

$$\begin{array}{lll} \text{minimizar} & x_1 - x_2 & \\ \text{sujeito a :} & x_1 + x_2 + x_3 & = 2 \\ & x_1 + x_4 & = 2 \\ & x_1,\ x_2,\ x_3,\ x_4 \geq 0. & \end{array}$$

O problema fase 1 é o PPL

$$\begin{array}{lll} \min & x_5 + x_6 & \\ \text{s. a :} & x_1 + x_2 + x_3 + x_5 & = 2 \\ & x_1 + x_4 + x_6 & = 2 \\ & x_1,\ x_2,\ x_3,\ x_4,\ x_5,\ x_6 \geq 0. & \end{array}$$

Seja a base ótima

$$B_* = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix},$$

para o conjunto de índices base $I_{B_*} = \{1, 5\}$, *com o vetor de variáveis básicas* $\hat{x}_{B_*} = [x_1, x_5]^T = [2, 0]^T$. *Observe que* x_5 *e* x_6 *são variáveis artificiais, e* x_5 *é uma variável básica. Pede-se: substituir* x_5 *por uma nova variável básica não artificial e fornecer* $I_{B_0},\ I_{N_0},\ B_0,\ N_0$ *e* x^0 *para a fase 2.*

Vejamos: consideremos a expressão obtida no Capítulo 5,

$$\hat{x}_{B_*} = B_*^{-1} b - B_*^{-1} N_* x_{N_*},$$

para a base ótima B_* *para o problema fase 1, tal que* $\hat{x}_{B_*}$ *é o vetor de variáveis básicas incluindo a variável artificial* x_5 *que é variável básica. Temos que*

$$N_* = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix},$$

para $I_{N_*} = \{2, 3, 4, 6\}$. *Ainda,* $x_{N_*} = \hat{x}_{N_*} = 0$ *para a matriz base* B_*. *A solução ótima para a fase 1,*

$$\hat{x} = [2, 0, 0, 0, 0, 0]^T,$$

é uma solução básica viável degenerada.

Agora, consideremos eliminadas as colunas referentes às variáveis artificiais na matriz não base N_**, isto é,*

$$N_* = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

para $I_{N_*} = \{2, 3, 4\}$*. Como* $x_{N_*} = 0$ *para* B_**, devemos substituir os índices de* I_{B_*}*, associados às variáveis artificiais, pelos índices de* I_{N_*}*. Uma maneira de fazer isto é verificar as linhas em* $B_*^{-1}N_*$*, correspondentes às variáveis artificiais em* $\hat{x}_{B_*}$*, que são não nulas. Estas linhas sempre existirão, porque a matriz* A *tem posto completo (veja página 171 em [11]). Em particular,*

$$B_*^{-1}N_* = \begin{bmatrix} 0 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 1 & -1 \end{bmatrix},$$

tal que a primeira linha corresponde a x_1 *e a segunda a* x_5*.*

A partir daí, troque as variáveis artificiais básicas em I_{B_*}*, verificando as linhas não nulas em* $B_*^{-1}N_*$*, pelas variáveis não básicas (não artificiais) em* I_{N_*}*, associadas às colunas não nulas em* $B_*^{-1}N_*$*; que são as mesmas associadas às colunas em* N_**. Em particular, a segunda linha em* $B_*^{-1}N_*$*, que corresponde a* x_5*, tem todos os coeficientes não nulos. Assim, basta tomar alguns deles, por exemplo, o primeiro coeficiente não nulo da segunda linha está na primeira coluna, que corresponde a* x_2*. Então,*

$$I_{B_*} = \{1, 2\} \quad e \quad I_{N_*} = \{5, 3, 4\}.$$

Finalmente, todas as variáveis artificiais restantes são variáveis não básicas. Então, basta eliminá-las como anteriormente no caso mais simples. Em particular,

$$I_{B_0} = \{1, 2\} \quad e \quad I_{N_0} = \{3, 4\},$$

com

$$B_0 = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \quad e \quad N_0 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Logo, a solução básica viável inicial para a fase 2 é dada por

$$x^0 = [2, 0, 0, 0]^T.$$

Atenção: Que tal fazer o exercício 4 agora?

A seguir, vamos enunciar um procedimento para a transição da fase 1 para a fase 2, especificando a troca das variáveis artificiais básicas em I_{B_*} pelas variáveis não básicas, não artificiais, em I_{N_*}.

Procedimento 11.1.5 *Fase 1 para a fase 2.*

Dados: uma matriz aumentada $\bar{A} = [A \;\; I]$, uma matriz base B_*, conjuntos de índices base I_{B_*}, não base I_{N_*} e $J = \{n+1, \ldots, n+m\}$, e uma solução ótima para (P_1) denotada por $z \in R^{n+m}$.

Se alguma variável artificial na solução ótima de (P_1) é variável básica,

$K := I_{B_*} \cap J$.

$I_{N_*} := I_{N_*} - (J - K)$.

Calcule a matriz D, $m \times |I_{N_*}|$,

$$B_* D = \bar{A}_{I_{N_*}}.$$

Para $i = 1 : m$

Se $(z_{B_*})_i$ é uma variável artificial,

Encontre $j = 1 : |I_{N_*}|$, tal que $D_{ij} \neq 0$.

Tome

$$i_b := (I_{N_*})_j; \quad i_{nb} := (I_{B_*})_i;$$

$$I_{B_*} := I_{B_*} \cup \{i_b\} - \{i_{nb}\}; \quad I_{N_*} := I_{N_*} - \{i_b\}.$$

$J := \emptyset$.

Saída:

$$I_B := I_{B_*}; \quad I_N := I_{N_*} - J; \quad B := B_*;$$

$$x_B := z_B, \quad x_N := 0 \text{ (0 é um vetor em } R^{n-m}\text{)}.$$

A saída deste procedimento fornece, para a próxima fase, uma matriz base inicial B, conjuntos de índices base I_B e não base I_N iniciais, e uma solução básica viável inicial $x \in R^n$ composta de um vetor de variáveis básicas x_B e de um vetor de variáveis não básicas x_N.

Agora estamos prontos para estudar a fase 2. É o que introduziremos no próximo capítulo.

Atenção: Que tal fazer o exercício 5 agora?

11.2 Exercícios

1. Considere o PPL

$$\begin{array}{lll} minimizar & -x_1 - 2x_2 & \\ sujeito\ a: & x_1 + x_2 + x_3 & = 4 \\ & 2x_1 + x_2 + x_4 & = 6 \\ & x_1 + x_5 & = 3 \\ & x_1,\ x_2,\ x_3,\ x_4,\ x_5 \geq 0. & \end{array}$$

   Pede-se:

   (a) Resolva graficamente este problema.
   (b) Quantas são as soluções básicas viáveis? E, quantos são os pontos extremos?
   (c) Por que, neste caso, temos mais soluções básicas viáveis do que pontos extremos?
   (d) Calcule o valor da função objetivo para cada solução básica viável. Quem é uma solução ótima?
   (e) Use o Algoritmo 11.1.1 fixando escolhas e atualizações, para resolver este problema.

2. Considere o PPL

$$\begin{array}{lll} minimizar & -x_1 & \\ sujeito\ a: & x_1 + x_2 & = 3 \\ & x_1,\ x_2 \geq 0. & \end{array}$$

   Pede-se:

   (a) Transforme este problema no problema fase 1.
   (b) Existe uma solução ótima para o problema fase 1? Justifique.

3. Considere o PPL do exercício anterior. Forneça uma solução ótima para o problema fase 1. Por outro lado, construa um problema inviável e certifique, através da fase 1, que de fato o seu problema é um problema inviável.

4. Considere o Exemplo 11.1.4. Substituindo x_5 por x_4, determine

$$I_{B_0}, \quad I_{N_0}, \quad B_0 \text{ e } x^0.$$

5. Implemente, em MATLAB ou em OCTAVE, a transição da fase 1 para a fase 2 para o problema de PL do Exemplo 11.1.4, com as entradas:

$$\bar{A} = \begin{bmatrix} 1 & 1 & 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 & 0 & 1 \end{bmatrix}, \quad B_* = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix},$$

$$I_{B_*} = \{1,\ 5\}, \quad I_{N_*} = \{2,\ 3,\ 4,\ 6\}, \quad J = \{5,\ 6\},$$

$$z = [2,0,0,0,0,0]^T, \quad m = 2 \text{ e } n = 4.$$

Capítulo 12

Método Simplex: algoritmo mestre adaptado

Continuamos com o nosso objetivo de enunciar e demonstrar que o método simplex fases 1 e 2 com a regra de Bland converge. Neste capítulo, estudamos um algoritmo mestre adaptado na tentativa de construir o algoritmo simplex à luz do Algoritmo 11.1.1.

12.1 Algoritmo mestre adaptado

Nesta seção, enunciamos um algoritmo mestre adaptado com um maior detalhamento para a próxima fase do método simplex.

12.1.1 Fase 2

Para a fase 2 consideramos o PPL (P).

Continuamos o nosso processo construtivo com uma nova pergunta. A partir da fase 1 concluída e, supondo que o problema original (P) não é um problema inviável e que já eliminamos todas as variáveis artificiais, partimos de uma solução básica viável e devemos escolher qual variável não básica passará a variável básica e qual variável básica passará a variável não básica. A este processo de escolha damos o nome de refinamento. E, este processo, é a idéia chave para o método simplex (veja página 103 em [9]). A propósito, a nossa segunda pergunta é:

como escolher novas variáveis básica e não básica?

Consideremos o problema original (P). Denotamos uma solução básica viável para (P), a saber, $\hat{x}$, associada a uma matriz base B. Denotamos, também, uma matriz não base N. Por definição,

$$\hat{x}_B = B^{-1}b \text{ e } \hat{x}_N = 0.$$

Por conveniência, particionamos a matriz tecnológica de posto completo A e o vetor custo c como

$$A = [N \quad B] \quad \text{e} \quad c = \begin{bmatrix} c_N \\ c_B \end{bmatrix},$$

onde B é uma matriz base $m \times m$ e N é uma matriz não base $m \times (n-m)$. Denotamos I_B o conjunto de índices base associados às variáveis básicas e I_N o conjunto de índices não base associados às variáveis não básicas.

Uma vez que uma matriz base é conhecida, todo ponto viável x para (P) pode ser rearranjado em uma ordem correspondente como

$$x = \begin{bmatrix} x_N \\ x_B \end{bmatrix}.$$

Desenvolvendo $Ax = b$,

$$\begin{aligned} Ax = b \Leftrightarrow Nx_N + Bx_B = b \Leftrightarrow Bx_B = b - Nx_N \Leftrightarrow \\ \Leftrightarrow x_B = B^{-1}b - B^{-1}Nx_N. \end{aligned} \tag{12.1}$$

Desenvolvendo $c^T x$ e usando a última igualdade,

$$\begin{aligned} c^T x = c_N^T x_N + c_B^T x_B = c_N^T x_N + c_B^T(B^{-1}b - B^{-1}Nx_N) = \\ = c_B^T B^{-1}b + (c_N^T - c_B^T B^{-1}N)x_N \Leftrightarrow \\ \Leftrightarrow c^T x = c_B^T B^{-1}b + \begin{bmatrix} c_N - (B^{-1}N)^T c_B \\ 0 \end{bmatrix}^T \begin{bmatrix} x_N \\ x_B \end{bmatrix}, \end{aligned} \tag{12.2}$$

onde $0 \in R^m$ denota um vetor de zeros.

Definição 12.1.1 *Dizemos que o vetor*

$$s = \begin{bmatrix} c_N - (B^{-1}N)^T c_B \\ 0 \end{bmatrix},$$

é denominado vetor custo reduzido.

Atenção: Que tal fazer o exercício 1 agora?

De acordo com a expressão (12.2), o vetor custo reduzido é o vetor das taxas de redução no valor da função objetivo com respeito à mudança na variável não básica.

Atenção: Que tal fazer o exercício 2 agora?

O próximo teorema fornece uma condição suficiente para uma solução básica viável ser uma solução ótima, cuja demonstração pode ser encontrada nas páginas 32 e 33 em [18].

Teorema 12.1.2 *Se $\hat{x}$ é uma solução básica viável com vetor custo reduzido não negativo, então $\hat{x}$ é uma solução ótima para o problema (P).*

Demonstração: Suponha $\hat{x}$ uma solução básica viável associada a uma matriz base B e o vetor custo reduzido $s \geq 0$. Considere x um ponto viável qualquer para o problema (P). Então, usando (12.2) e a definição de vetor custo reduzido, obtemos

$$c^T x - c^T \hat{x} = c_B^T B^{-1} b + s^T \begin{bmatrix} x_N \\ x_B \end{bmatrix} - c_B^T B^{-1} b = s^T \begin{bmatrix} x_N \\ x_B \end{bmatrix} \geq 0.$$

Portanto, $c^T \hat{x} \leq c^T x$, para todo x viável. Isto significa que $\hat{x}$ é uma solução ótima, finalizando a demonstração. ∎

Atenção: Que tal fazer o exercício 3 agora?

Examinando a literatura em Programação Linear, constatamos que a notação para o desenvolvimento do método simplex é fundamental para uma exposição clara e fácil. Neste ponto, a idéia é excluir a matriz N na implementação do algoritmo simplex (se necessário, utilizaremos N uma única vez no Procedimento 11.1.5) e trabalhar com os conjuntos de índices base e não base.

Notação: Referindo-nos à matriz A particionada pela matriz base B e pela matriz não base N, para $j_l \in \{1, \ldots, n-m\}$, $l \in I_N$, $i \in \{1, \ldots, m\}$ e $k_i \in I_B$, denotamos $d(l) \in R^m$, a solução única do sistema

$$Bd(l) = N_{j_l} = A_l.$$

Finalmente, denotamos

$$d_i(l) \in R,$$

a coordenada i do vetor $d(l)$ associada à coluna k_i da matriz A.

Atenção: Que tal fazer o exercício 4 agora?

O próximo resultado certifica quando um PPL é um problema ilimitado. Esta demonstração pode ser encontrada na página 135 em [11].

Teorema 12.1.3 *Considere o PPL (P). Seja dada uma solução básica viável $\hat{x}$ associada a uma matriz base B. Se $s_h < 0$ e $d(h) \leq 0$, para algum $h \in I_N$, então (P) é um problema ilimitado.*

Demonstração: Partimos de uma solução básica viável $\hat{x}$. Fixe $h \in I_N$. Seja $\hat{x}_h = 0$ uma variável não básica de $\hat{x}$ tal que $s_h < 0$ e $d(h) \leq 0$. A partir de $\hat{x}$, tentaremos encontrar uma nova solução básica viável x, fazendo x_h a nova variável básica, isto é, atribuindo-lhe um valor positivo. As demais variáveis não básicas de $\hat{x}$ continuarão nulas, isto é,

$$x_l = \hat{x}_l = 0, \text{ para todo } l \in I_N, \ l \neq h. \tag{12.3}$$

Substituindo (12.3) em (12.1), temos:

$$x_B = \hat{x}_B - d(h)x_h.$$

Substituindo (12.3) em (12.2), temos:

$$c^T x = c^T \hat{x} + s_h x_h.$$

Como $d(h) \leq 0$, podemos fazer x_h crescer tanto quanto queiramos sem o risco de alguma componente de x assumir valores negativos. Fazendo $x_h \to \infty$, temos, devido a última igualdade, $c^T x \to -\infty$, uma vez que $s_h < 0$. Logo, (P) é um problema ilimitado. Isto finaliza a demonstração. ∎

Atenção: Que tal fazer o exercício 5 agora?

O próximo resultado refere-se a bases de um espaço vetorial, o qual será importante para a demonstração do próximo teorema. Sua demonstração pode ser encontrada nas páginas 36 e 37 em [11].

Proposição 12.1.4 *Seja $\mathcal{B} = \{u^1, \ldots, u^m\}$ uma base do espaço vetorial $\mathcal{V}$ e seja o vetor $v \in \mathcal{V}$, que pode ser escrito da forma $v = w_1 u^1 + \ldots + w_m u^m$. Caso exista algum t, $t = 1, \ldots, m$, tal que $w_t \neq 0$, então também o conjunto*

$$\mathcal{B}' = \{u^1, \ldots, u^{t-1}, v, u^{t+1}, \ldots, u^m\}$$

será uma base de $\mathcal{V}$.

Demonstração: Suponhamos, sem perda de generalidade, que $t = 1$. Qualquer vetor $z \in \mathcal{V}$ pode ser escrito como combinação linear dos vetores da base $\mathcal{B}$, isto é, $z = z_1 u^1 + \ldots + z_m u^m$. Como $w_1 \neq 0$, podemos escrever também

$$v = w_1 u^1 + \ldots + w_m u^m \Rightarrow u^1 = \frac{1}{w_1}(v - w_2 u^2 - \ldots - w_m u^m).$$

Substituindo u^1 na primeira igualdade em z, temos:

$$z = z_1[\frac{1}{w_1}(v - w_2 u^2 - \ldots - w_m u^m)] + z_2 u^2 + \ldots + z_m u^m,$$

isto é,

$$z = \frac{z_1}{w_1} v + (z_2 - \frac{z_1 w_2}{w_1}) u^2 + \ldots + (z_m - \frac{z_1 w_m}{w_1}) u^m.$$

Assim, qualquer vetor $z \in \mathcal{V}$ pode ser escrito como combinação linear dos vetores $v, u^2, \ldots, u^m$. Por outro lado, fazendo

$$\lambda_1 v + \lambda_2 u^2 + \ldots + \lambda_m u^m = 0,$$

segue-se pela substituição de v como combinação linear de $u^1, \ldots, u^m$ que

$$\lambda_1 w_1 u^1 + (\lambda_1 w_2 + \lambda_2) u^2 + \ldots + (\lambda_1 w_m + \lambda_m) u^m = 0.$$

Como os vetores $u^1, \ldots, u^m$ são linearmente independentes (*li*),

$$\lambda_1 w_1 = (\lambda_1 w_2 + \lambda_2) = \ldots = (\lambda_1 w_m + \lambda_m) = 0.$$

Como $w_1 \neq 0$, temos $\lambda_1 = 0$. Segue-se pelas demais equações que

$$\lambda_1 = \ldots = \lambda_m = 0,$$

o que significa que os vetores $v, u^2, \ldots, u^m$ são *li*. Então, pela definição de base de um espaço vetorial,

$$\mathcal{B}' = \{v, u^2, \ldots, u^m\}$$

é uma base de $\mathcal{V}$, porque os vetores $v, u^2, \ldots, u^m$ geram o espaço vetorial $\mathcal{V}$ e são *li*. Pela arbitrariedade de $t = 1$, concluímos a demonstração. ∎

Esta proposição garante que para uma base $\mathcal{B}$ de um espaço vetorial $\mathcal{V}$, podemos substituir um vetor da base por um outro vetor obtendo novamente uma outra base de $\mathcal{V}$, verificada uma condição bastante simples.

Atenção: Que tal fazer o exercício 6 agora?

O teorema a seguir fornece um critério de possível melhoria para o valor da função objetivo do problema (P), cuja demonstração pode ser encontrada nas páginas 136 e 137 em [11].

Teorema 12.1.5 *Considere o PPL (P). Seja dada uma solução básica viável $\hat{x}$ associada a uma matriz base B, e uma matriz não base N. Considere $s_h < 0$, para algum $h \in I_N$, tal que existe $d_i(h) > 0$ ao menos para algum $k_i \in I_B$. Ainda, considere*

$$\frac{\hat{x}_{k_q}}{d_q(h)} = \min_{i=1,\ldots,m} \{\frac{\hat{x}_{k_i}}{d_i(h)};\ d_i(h) > 0,\ k_i \in I_B\}. \tag{12.4}$$

Então, fazendo $x_h = \hat{x}_{k_q}/d_q(h)$ a nova variável básica, anulamos x_{k_q} fazendo-a variável não básica, obtendo assim uma nova solução básica viável x tal que $c^T x \leq c^T \hat{x}$, com desigualdade estrita caso o valor x_h seja positivo.

Demonstração: Considere $i = 1, \ldots, m$. Seja x a nova solução obtida a partir da solução básica viável $\hat{x}$. Vamos demonstrar que x é também solução básica viável. Na nova solução, mantemos todas as variáveis não básicas, exceto x_h, isto é,

$$x_l = \hat{x}_l = 0,\ \text{para todo}\ l \in I_N,\ l \neq h. \tag{12.5}$$

Fazemos

$$x_h = \frac{\hat{x}_{k_q}}{d_q(h)} \geq 0.$$

Substituindo (12.5) e x_h em (12.1), temos, para as variáveis básicas de $\hat{x}$, os novos valores:

$$x_{k_i} = \hat{x}_{k_i} - d_i(h)x_h = \hat{x}_{k_i} - d_i(h)\frac{\hat{x}_{k_q}}{d_q(h)}, \text{ para todo } k_i \in I_B. \tag{12.6}$$

Suponha satisfeita a condição (12.4), isto é,

$$\frac{\hat{x}_{k_i}}{d_i(h)} \geq \frac{\hat{x}_{k_q}}{d_q(h)}$$

para todo $k_i \in I_B$ com $d_i(h) > 0$. Logo,

$$x_{k_i} = \hat{x}_{k_i} - d_i(h)\frac{\hat{x}_{k_q}}{d_q(h)} \geq \hat{x}_{k_i} - d_i(h)\frac{\hat{x}_{k_i}}{d_i(h)} = 0,$$

para todo $k_i \in I_B$ com $d_i(h) > 0$.

Para $d_i(h) \leq 0$ temos, pela definição de $x_h \geq 0$ e por (12.6), $x_{k_i} \geq 0$. Logo, $x_{k_i} \geq 0$ para todo $k_i \in I_B$. Em particular, temos por (12.6) $x_{k_q} = 0$. Considerando x_{k_q} como variável não básica e fazendo x_h a nova variável básica, verificamos ter m variáveis básicas não negativas assim como $n - m$ variáveis não básicas nulas. Temos, portanto, satisfeitas as condições de não negatividade: $x \geq 0$. Por outro lado, como todas as variáveis satisfazem o sistema de equações (12.1) e este é equivalente ao sistema $Ax = b$, temos satisfeitas todas as restrições do PPL (P).

Para mostrar que temos uma solução básica viável, basta mostrar que podemos associar uma matriz base às variáveis básicas. De acordo com a definição de $d(h)$, podemos escrever

$$A_h = \sum_{k_i \in I_B} d_i(h) A_{k_i}.$$

Como $d_i(h) > 0$, para algum $k_i \in I_B$, podemos de acordo com a Proposição 12.1.4, ter uma nova matriz base trocando na matriz base antiga o vetor A_{k_q} por A_h. Temos, portanto, uma solução básica viável x.

Usando (12.2), (12.5) e a definição de $x_h \geq 0$,

$$c^T x - c^T \hat{x} = s_h x_h \leq 0,$$

uma vez que $s_h < 0$. Aqui $c^T x < c^T \hat{x}$ se $x_h > 0$. Isto finaliza a demonstração.
∎

Atenção: Que tal fazer o exercício 7 agora?

Neste ponto, podemos responder a nossa segunda pergunta de como escolher novas variáveis básica e não básica. De fato, este último teorema fornece a resposta: basta escolher algum índice $l \in I_N$ com $s_l < 0$, tal que x_l é a nova variável básica, e escolher algum índice $k_i \in I_B$ satisfazendo (12.4), tal que x_{k_i} é a nova variável não básica.

Como se trata de escolhas, a seguir vamos enunciar o algoritmo mestre (Algoritmo 11.1.1) de maneira, digamos, mais completa. Relembramos que, fixado um índice j, denotamos A_j a j-ésima coluna de A.

Algoritmo 12.1.6 *Mestre Adaptado.*

Dados: uma solução básica viável x^0 associada a uma matriz base inicial B_0, um conjunto de índices base I_{B_0} e um conjunto de índices não base I_{N_0}.

$k := 0.$

REPITA

Calcule o vetor multiplicador simplex $y \in R^m$,

$$B_k^T y = c_{B_k}.$$

Calcule o vetor custo reduzido $s \in R^n$ tal que,

$$s_l = 0, \text{ para todo } l \in I_{B_k} \text{ e}$$

$$s_l = c_l - y^T A_l, \text{ para todo } l \in I_{N_k}.$$

Se $s \geq 0$, então PARE; solução ótima x^k.

Entrada na base: escolha $h \in I_{N_k}$ tal que $s_h < 0$.

Calcule o vetor (direção) $d(h) \in R^m$,

$$B_k d(h) = A_h.$$

Se $d(h) \leq 0$, então PARE; problema ilimitado.

Saída da base: escolha $k_q \in I_{B_k}$ tal que

$$\frac{x_{k_q}^k}{d_q(h)} = \min_{i=1,\ldots,m} \{\frac{x_{k_i}^k}{d_i(h)}; \quad d_i(h) > 0, \quad k_i \in I_{B_k}\}.$$

Atualize os índices base e não base, respectivamente,

$$I_{B_{k+1}} := I_{B_k} \cup \{h\} - \{k_q\};$$

$$I_{N_{k+1}} := I_{N_k} \cup \{k_q\} - \{h\}.$$

Atualize a matriz base

$$B_{k+1} := B_k + (A_h - A_{k_q})e_q^T,$$

onde $e_q \in R^m$ *é um vetor de zeros com valor um na posição* $q = 1, \ldots, m$.

Novo ponto: calcule a nova solução básica viável $x^{k+1} \in R^n$,

$$B_{k+1}x_{B_{k+1}}^{k+1} = b;$$

$$x_{k_q}^{k+1} := 0.$$

$k := k + 1.$

ATÉ QUE 'convirja'.

Temos algumas observações a fazer acerca deste algoritmo. Inicialmente, observe que são dados os conjuntos de índices base I_{B_0} e não base I_{N_0}. Isto se deve à nossa conveniência de escrita e implementação. Depois, o vetor custo reduzido s é calculado pela Definição 12.1.1 em dois passos: primeiro, referenciamos o vetor multiplicador simplex y, o qual se relaciona com a variável dual; e segundo, o cálculo das coordenadas de s associadas ao conjunto de índices não base da iteração corrente é feito usando o vetor multiplicador simplex ao invés do cálculo de inversão de matrizes, conforme a Definição 12.1.1.

Também, observe que os critérios de parada exibindo uma solução ótima é devido ao Teorema 12.1.2 e, certificando problema ilimitado, é devido ao Teorema 12.1.3. Ainda, quando possíveis, tanto as escolhas para a entrada na base quanto para a saída da base são devidas ao Teorema 12.1.5.

Finalmente, observe que elaboramos uma terceira pergunta:

como atualizar matriz base e solução básica viável?

A resposta foi imediata: da maneira que se encontra no Algoritmo 12.1.6. Isto é, trocando na matriz base a coluna A_{k_q} pela coluna A_h e resolvendo o sistema de equações lineares $Bx_B = b$. A propósito, note que as variáveis não básicas são nulas. Daí, devemos tomar a nova variável não básica x_{k_q} igual a zero. Além disso, atualizamos os conjuntos de índices base e não base.

Atenção: Que tal fazer o exercício 8 agora?

Neste ponto, gostaríamos de afirmar que o Algoritmo 12.1.6, define a nossa estratégia para o desenvolvimento do método simplex (revisado) fase 1 e fase 2.

No próximo capítulo, estudaremos a convergência deste algoritmo, através de um exemplo, em que fixaremos uma certa escolha para a entrada na base e para a saída da base.

Atenção: Que tal fazer o exercício 9 agora?

12.2 Exercícios

1. Considere o PPL

$$\begin{array}{lll} \textit{minimizar} & -x_1 - 2x_2 & \\ \textit{sujeito a:} & x_1 + 4x_2 + x_3 & = 4 \\ & x_1 - x_2 + x_4 & = 3 \\ & x_1,\ x_2,\ x_3,\ x_4 \geq 0. & \end{array}$$

   Pede-se:

   (a) Defina a matriz A e os vetores b e c.

   (b) Forneça uma matriz base B e uma matriz não base N.

   (c) Para B e N defina, respectivamente, o conjunto de índices base I_B e não base I_N.

   (d) Para esta matriz base B, calcule o vetor custo reduzido.

2. Calcule o vetor custo reduzido do PPL do exercício anterior para a matriz base dada pela matriz identidade. Verifique que a componente mais negativa do vetor custo reduzido induz um hiperplano no conjunto viável com valor da função objetivo igual a -2. Todavia, verifique que

poderíamos ter obtido um novo ponto extremo adjacente, x, com maior redução do valor da função objetivo, isto é, $c^T x = -3$.

3. Calcule o vetor custo reduzido do PPL do exercício anterior para a matriz base B obtida pelas primeira e segunda colunas de A. O que você conclui?

4. Seja o sistema de equações lineares

$$\begin{aligned} x_1 + x_2 + x_3 \quad &= 4 \\ 2x_1 + x_2 + x_4 \quad &= 6 \\ x_1 + x_5 \qquad &= 3. \end{aligned}$$

Pede-se:

(a) Para

$j_l \in \{1,\ 2\},\ l \in I_N = \{3,\ 4\},\ i \in \{1,\ 2,\ 3\}$ e $k_i \in I_B = \{1,\ 2,\ 5\}$,

verifique que fixados $j_l = 1$ e $l = 3$ (ou $j_l = 2$ e $l = 4$),

$$Bd(l) = N_{j_l}$$

é o mesmo que

$$Bd(l) = A_l.$$

(b) Calcule $d(h)$ e $d_q(h)$, para $l = h = 4$ e $k_i = k_q = k_3 = 5$.

5. Considere o PPL

$$\begin{array}{lcl} \textit{minimizar} & -x_1 & \\ \textit{sujeito a}: & x_1 - x_2 & = 2 \\ & x_1,\ x_2 \geq 0. & \end{array}$$

Pede-se:

(a) Resolva graficamente este problema.

(b) Para o conjunto de índices base $I_B = \{1\}$ e a solução básica viável $\hat{x} = [2,\ 0]^T$, use o Teorema 12.1.3 para concluir que este PPL é um problema ilimitado.

6. Considere o sistema

$$\begin{aligned} x_1 + x_2 + x_3 &= 4 \\ 2x_1 + x_2 + x_4 &= 6. \end{aligned}$$

   Para a matriz

$$A = \begin{bmatrix} 1 & 1 & 1 & 0 \\ 2 & 1 & 0 & 1 \end{bmatrix},$$

   defina a matriz base

$$B = \begin{bmatrix} 1 & 1 \\ 2 & 0 \end{bmatrix},$$

   e tome $v = [1, 1]^T$. O vetor $[1, 2]^T$ (ou $[1, 0]^T$) pode ser substituído pelo vetor v? Por quê?

7. Considere o PPL

$$\begin{array}{ll} \textit{minimizar} & x_1 \\ \textit{sujeito a :} & x_1 + x_2 = 2 \\ & x_1,\ x_2 \geq 0. \end{array}$$

   Pede-se:

   (a) Resolva graficamente este problema.

   (b) Para a solução básica viável $\hat{x} = [2,\ 0]^T$, use o Teorema 12.1.5 para concluir que a nova variável x possui $c^T x \leq c^T \hat{x}$. Quem é x?

8. Implemente, em MATLAB ou em OCTAVE, a fase 2 para o PPL do exercício 1 e para o PPL do exercício 7, através do Algoritmo 12.1.6. Observe que este algoritmo só é implementável a partir de escolhas para a entrada na base e para a saída da base. Assim, tome decisões para a fase 1 e tome decisões para as escolhas (o chamado refinamento). Acredite em você, garota(o)!

9. Faça uma pesquisa (biblioteca é um lugar legal!) da literatura em PL e compare o método simplex com o método simplex revisado.

Capítulo 13

Método Simplex: algoritmo fases 1 e 2

Neste capítulo concluímos o enunciado e a convergência do algoritmo simplex de duas fases com a regra de Bland.

13.1 Algoritmo fases 1 e 2

Consideremos o PPL (P). Nesta seção, enunciamos o algoritmo de duas fases, também, de uma maneira construtiva.

Neste ponto, retiramos o apóstrofo do termo 'ingênua', usado no início do Capítulo 11, quanto à idéia do método simplex. Isto se dá devido ao fenômeno denominado ciclagem, que pode ocorrer aplicando o método simplex, quando existe uma solução básica viável degenerada. A ciclagem consiste em voltarmos para a mesma matriz base depois de um certo número de iterações. Desta forma, o método simplex pode gerar uma seqüência divergente.

Muitos autores afirmaram que a ciclagem dificilmente ocorria em problemas práticos. Todavia, Goldbarg (página 118 em [19]) deparou-se com vários exemplos de ciclagem na solução de problemas de particionamento não ponderado de grande porte.

Beale [7] construiu um exemplo de PPL que cicla, com a utilização do método simplex, para uma escolha particular para a entrada na base. A seguir, executaremos o Algoritmo 12.1.6 para o exemplo de Beale (conforme página 274 em [7]), no formato padrão, escolhendo para a entrada na base o índice base com menor custo reduzido. Para a saída da base, escolhemos o menor índice, quando houver empate para a saída.

Exemplo 13.1.1 *Verifique ciclagem no exemplo de Beale, a saber:*

$$\begin{array}{lll} \textit{minimizar} & -\frac{3}{4}x_1 + 20x_2 - \frac{1}{2}x_3 + 6x_4 & \\ \textit{sujeito a}: & \frac{1}{4}x_1 - 8x_2 - x_3 + 9x_4 + x_5 & = 0 \\ & \frac{1}{2}x_1 - 12x_2 - \frac{1}{2}x_3 + 3x_4 + x_6 & = 0 \\ & x_3 + x_7 & = 1 \\ & x_1,\ x_2,\ x_3,\ x_4,\ x_5,\ x_6,\ x_7 \geq 0. & \end{array}$$

Vejamos: a matriz A e os vetores b e c são

$$A = \begin{bmatrix} \frac{1}{4} & -8 & -1 & 9 & 1 & 0 & 0 \\ \frac{1}{2} & -12 & -\frac{1}{2} & 3 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 \end{bmatrix},\ b = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}\ e\ c = \begin{bmatrix} -\frac{3}{4} \\ 20 \\ -\frac{1}{2} \\ 6 \\ 0 \\ 0 \\ 0 \end{bmatrix}.$$

Aplicando o Algoritmo 12.1.6, temos dados a matriz base inicial, o ponto inicial, os conjuntos de índices base e não base iniciais, respectivamente,

$$B_0 = I,\quad x^0 = [0\ 0\ 0\ 0\ 0\ 0\ 1]^T,\quad I_{B_0} = \{5,\ 6,\ 7\}\ e\ I_{N_0} = \{1,\ 2,\ 3,\ 4\}.$$

Calculando o vetor multiplicador simplex,

$$B_0^T y = c_{B_0} \Rightarrow Iy = 0 \Rightarrow y = 0.$$

Usamos $y = 0$ para calcular o vetor custo reduzido, para $l \in I_{N_0}$,

$$s_1 = c_1 - y^T A_1 = -3/4,\quad s_2 = c_2 - y^T A_2 = 20,$$

$$s_3 = c_3 - y^T A_3 = -1/2\ e\ s_4 = c_4 - y^T A_4 = 6.$$

Fixamos a escolha do índice com menor custo reduzido para a entrada na base, logo, escolhemos $h = 1$. Calculando o vetor direção,

$$B_0 d(h) = A_h \Rightarrow Id(1) = A_1 \Rightarrow d(1) = [1/4\ \ 1/2\ \ 0]^T.$$

Fixamos, para a saída da base, a escolha do menor índice tal que

$$\min_{i=1,2,3}\{\frac{x^0_{k_i}}{d_i(h)};\quad d_i(h)>0,\quad k_i\in I_{B_0}\} = \min\{\frac{x^0_{k_1}}{d_1(1)},\ \frac{x^0_{k_2}}{d_2(1)}\} = 0.$$

Ou seja, $k_q = k_1 = 5$*. Para atualizar a nova solução básica viável, temos*

$$I_{B_1} = \{1,\ 6,\ 7\}\ e\ I_{N_1} = \{5,\ 2,\ 3,\ 4\},$$

$$B_1 = \begin{bmatrix} \frac{1}{4} & 0 & 0 \\ \frac{1}{2} & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

logo,

$$B_1 x^1_{B_1} = b \Rightarrow \begin{bmatrix} \frac{1}{4} & 0 & 0 \\ \frac{1}{2} & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x^1_1 \\ x^1_6 \\ x^1_7 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \Rightarrow x^1_{B_1} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix};$$

além disso,

$$x^1_{k_q} = x^1_{k_1} = x^1_5 = 0.$$

Então, $x^1 = [0\ 0\ 0\ 0\ 0\ 0\ 1]^T$.

Para as próximas iterações temos:

$$I_{B_2} = \{1,\ 2,\ 7\},\quad I_{B_3} = \{3,\ 2,\ 7\},\quad I_{B_4} = \{3,\ 4,\ 7\},$$

$$I_{B_5} = \{5,\ 4,\ 7\}\ e\ I_{B_6} = \{5,\ 6,\ 7\} = I_{B_0},$$

para $x^0 = x^1 = \ldots = x^6$*. Isto estabelece um ciclo. Quer dizer,*

$$I_{B_7} = I_{B_1},\ I_{B_8} = I_{B_2},\ \ldots,\ I_{B_{12}} = I_{B_0},\ etc.,$$

e $x^0 = x^1 = \ldots = x^{12} = \ldots$*. Todavia, a solução ótima para este problema de Beale é o ponto*

$$x^* = [1\ 0\ 1\ 0\ 3/4\ 0\ 0]^T,\ com\ I_{B_*} = \{1,\ 3,\ 5\}.$$

Atenção: Que tal fazer o exercício 1 agora?

Existem algumas maneiras de escolher uma nova variável básica, isto é, uma nova coluna na matriz não base N para compor a matriz base B. E, também, existem algumas maneiras de escolher uma nova variável não básica, isto é, uma nova coluna em B para compor a matriz N. Todavia, escolheremos uma mesma regra que decide quem entra e quem sai da base. Optamos por uma escolha objetivando convergência do método simplex, como veremos adiante e, também, simplicidade de idéia e de implementação. Esta escolha é conhecida como a regra do menor índice ou a regra de Bland [9], a saber:

(i) entre todos os candidatos a entrar na base, selecione a variável x_h tendo o menor índice, isto é, encontre

$$h = \min_{l \in I_N} \{l;\ s_l < 0\};$$

(ii) entre todos os candidatos a sair da base, selecione a variável x_{k_q} tendo o menor índice, isto é, encontre

$$k_q = \min_{k_{i_0} \in I_{B_k}} \{k_{i_0};\ \frac{x^k_{k_{i_0}}}{d_{i_0}(h)} = \min_{i=1,\ldots,m} \{\frac{x^k_{k_i}}{d_i(h)};\ d_i(h) > 0,\ k_i \in I_{B_k}\}\}.$$

O elemento $d_q(h)$ é denominado pivô. Para o formato original (tabular) do método simplex, a q-ésima linha é denominada linha pivô e a h-ésima coluna é denominada coluna pivô.

Finalmente, estabeleceremos o algoritmo simplex revisado fases 1 e 2 para o PPL (P) com uma técnica anti-ciclagem, ou seja, com a regra de Bland. Se necessário, usaremos o Procedimento 11.1.5 para a transição da fase 1 para a fase 2.

Algoritmo 13.1.2 *Simplex Fases 1 e 2.*

Fase 1: obtemos uma solução básica viável x^0 associada a uma matriz base inicial B_0, um conjunto de índices base I_{B_0} e um conjunto de índices não base I_{N_0}, eliminando todas as variáveis artificiais. Caso contrário, certificamos problema inviável.

$k := 0$.

REPITA

Calcule o vetor multiplicador simplex $y \in R^m$,

$$B_k^T y = c_{B_k}.$$

Calcule o vetor custo reduzido $s \in R^n$ *tal que,*

$$s_l = 0, \text{ para todo } l \in I_{B_k} \text{ e}$$

$$s_l = c_l - y^T A_l, \text{ para todo } l \in I_{N_k}.$$

Se $s \geq 0$
então x^k *é uma solução ótima;*
senão
Entrada na base: calcule o novo índice base

$$h = \min_{l \in I_{N_k}} \{l;\ s_l < 0\}.$$

Calcule o vetor (direção) $d(h) \in R^m$,

$$B_k d(h) = A_h.$$

Se $d(h) \leq 0$
então problema ilimitado;
senão
Saída da base: calcule o novo índice não base

$$k_q = \min_{k_{i_0} \in I_{B_k}} \{k_{i_0};\ \frac{x^k_{k_{i_0}}}{d_{i_0}(h)} = \min_{i=1,\ldots,m} \{\frac{x^k_{k_i}}{d_i(h)};\ d_i(h) > 0,\ k_i \in I_{B_k}\}\}.$$

Atualize os índices base e não base, respectivamente,

$$I_{B_{k+1}} := I_{B_k} \cup \{h\} - \{k_q\};$$

$$I_{N_{k+1}} := I_{N_k} \cup \{k_q\} - \{h\}.$$

Atualize a matriz base

$$B_{k+1} := B_k + (A_h - A_{k_q})e_q^T,$$

onde $e_q \in R^m$ *é um vetor de zeros com valor um na posição* $q = 1, \dots, m$.

Novo ponto: calcule a nova solução básica viável $x^{k+1} \in R^n$,

$$B_{k+1}x^{k+1}_{B_{k+1}} = b;$$

$$x^{k+1}_{k_q} := 0.$$

$k := k+1$.

ATÉ QUE $(s \geq 0)$ *ou* $(d(h) \leq 0)$.

Atenção: Que tal fazer os exercícios 2, 3 e 4 agora?

O próximo resultado garante convergência para o algoritmo simplex. Sua demonstração pode ser encontrada em Bland [9], página 104 (veja páginas 46 e 47 em [37]).

Teorema 13.1.3 *O algoritmo simplex revisado fase 1 e fase 2, com a regra de Bland, converge.*

Demonstração: Sem perda de generalidade considere a fase 2 do algoritmo. Convergência significa verificar otimalidade ou detectar problema ilimitado. Suponha por contradição que o algoritmo simplex não converge, isto é, ocorre ciclagem. Uma vez que o algoritmo simplex determina unicamente um elemento pivô, o ciclo é bem determinado.

Seja $T \subset \{1, \dots, n\}$ o subconjunto de índices de todas as variáveis que entram na base durante o ciclo, isto é, $j \notin T$ significa que x_j jamais será ou, sempre será, uma variável básica durante a ciclagem. Considere

$$t = \max\{j;\ j \in T\}.$$

Considere uma iteração $\bar{k}$, com matriz base $\bar{B}$ e solução básica viável $\bar{x}$, tal que A_t sai da base e A_r entra na base. E, considere uma iteração $\hat{k}$, com matriz base $\hat{B}$ e solução básica viável $\hat{x}$, tal que A_t retorna para a base. Logo, $r \in T$. E, na ciclagem, o valor da função objetivo permanece o mesmo.

Na iteração $\bar{k}$ denotamos:

$$x_{\bar{B}} = \bar{x}_{\bar{B}} - \bar{B}^{-1}\bar{N}x_{\bar{N}}$$

e

$$c^T x = c^T \bar{x} + \bar{s}_{\bar{N}}^T x_{\bar{N}}; \tag{13.1}$$

e na iteração $\hat{k}$ denotamos:

$$x_{\hat{B}} = \hat{x}_{\hat{B}} - \hat{B}^{-1}\hat{N}x_{\hat{N}}$$

e

$$c^T x = c^T \hat{x} + \hat{s}_{\hat{N}}^T x_{\hat{N}}.$$

Na ciclagem temos que $c^T\hat{x} = c^T\bar{x}$ e, pelo algoritmo simplex temos que $\hat{s}_{\hat{B}} = 0$. Daí,

$$c^T x = c^T \hat{x} + \hat{s}_{\hat{N}}^T x_{\hat{N}} = c^T \bar{x} + \hat{s}^T x,$$

que deve ser verificada para

$$x_r = \lambda, \quad x_j = 0, \quad j \in I_{\bar{N}} - \{r\}$$

e

$$x_{k_i} = \bar{x}_{k_i} - \bar{d}_i(r)\lambda, \quad k_i \in I_{\bar{B}}, \quad i = 1, \ldots, m,$$

fornecendo, de (13.1),

$$c^T x = c^T \bar{x} + \bar{s}_r \lambda. \tag{13.2}$$

Então,

$$c^T x = c^T \bar{x} + \hat{s}^T x = c^T \bar{x} + \hat{s}_r \lambda + \sum_{i=1}^{m} \hat{s}_{k_i}(\bar{x}_{k_i} - \bar{d}_i(r)\lambda). \tag{13.3}$$

De (13.2) e (13.3),

$$c^T \bar{x} + \bar{s}_r \lambda = c^T \bar{x} + \hat{s}_r \lambda + \sum_{i=1}^{m} \hat{s}_{k_i}(\bar{x}_{k_i} - \bar{d}_i(r)\lambda),$$

e isolando os termos em λ,

$$(\bar{s}_r - \hat{s}_r + \sum_{i=1}^{m} \hat{s}_{k_i} \bar{d}_i(r))\lambda = \sum_{i=1}^{m} \hat{s}_{k_i} \bar{x}_{k_i},$$

cujo lado direito desta última igualdade é uma constante para qualquer λ. Logo, esta constante deve ser nula. Daí,

$$\bar{s}_r - \hat{s}_r + \sum_{i=1}^{m} \hat{s}_{k_i} \bar{d}_i(r) = 0. \tag{13.4}$$

Agora, o nosso raciocínio para o restante desta demonstração se concentrará nesta última igualdade. Assim, como A_r entra na base na iteração $\bar{k}$, $\bar{s}_r < 0$. E, como A_r não está entrando na base na iteração $\hat{k}$, então $\hat{s}_r \geq 0$, porque pela definição de t e pelo fato de que $r \in T$, concluímos que $r < t$ e, também, porque estamos usando a regra de Bland no algoritmo. Ainda, para que a igualdade (13.4) seja verificada, deve existir α, $\alpha = 1, \ldots, m$, tal que $\hat{s}_{k_\alpha} \bar{d}_\alpha(r) > 0$. Uma vez que $\hat{s}_{k_\alpha} \neq 0$, A_{k_α} está na base na iteração $\bar{k}$ e não está na base na iteração $\hat{k}$. Logo, $k_\alpha \in T$ e $k_\alpha \leq t$, para $k_\alpha \in I_{\bar{B}}$.

Como A_t sai da base na iteração $\bar{k}$ e A_r entra, então $\bar{d}_i(r) > 0$ para $k_i = t$, $i = 1, \ldots, m$. Como A_t retorna à base na iteração $\hat{k}$, $\hat{s}_t < 0$, logo, $\hat{s}_t \bar{d}_i(r) < 0$ para $k_i = t$, $i = 1, \ldots, m$. Segue-se que $k_\alpha < t$, para $k_\alpha \in I_{\bar{B}}$.

Para $k_\alpha \in I_{\bar{B}}$, se $k_\alpha < t$, então $\hat{s}_{k_\alpha} > 0$ ($\hat{s}_{k_\alpha} \neq 0$) implicando em $\bar{d}_\alpha(r) > 0$ para que possamos satisfazer $\hat{s}_{k_\alpha} \bar{d}_\alpha(r) > 0$. Todas as iterações de $\bar{k}$ e $\hat{k}$ estão associadas a soluções básicas viáveis degeneradas implicando que o valor de x_{k_α} seja o mesmo na base ou fora, isto é, igual a zero. Desta forma, $\bar{x}_{k_\alpha} = 0$ e, como $\bar{d}_\alpha(r) > 0$, então pela regra de Bland A_{k_α} deveria deixar a base na iteração $\bar{k}$, pois $k_\alpha < t$; contradizendo o fato de que A_t sai da base na iteração $\bar{k}$. Isto finaliza a demonstração. ∎

Podemos concluir, portanto, que o método simplex gera soluções básicas viáveis (degeneradas, inclusive) pela mudança de uma única coluna da matriz base em cada iteração, com o valor da função objetivo menor ou igual ao anterior. Usando técnicas anti-ciclagem (existem outras além da regra de Bland!), o método simplex não repete bases. Como o número de bases possíveis é finito, o algoritmo simplex converge.

Para uma implementação prática do método simplex, consulte por exemplo, Goldfarb e Reid [20].

No próximo capítulo vamos estudar o método afim-escala, através de um algoritmo mestre.

Atenção: Que tal fazer o exercício 5 agora?

13.2 Exercícios

1. Continue o Exemplo 13.1.1, fazendo as contas, para mais cinco iterações. Isto é, verifique que o algoritmo simplex pode não convergir se não refinarmos as nossas escolhas.

2. Implemente, em MATLAB ou em OCTAVE, o algoritmo simplex revisado, com a regra de Bland, para o problema de Beale.

3. Implemente, em MATLAB ou em OCTAVE, o algoritmo simplex revisado fases 1 e 2, com a regra de Bland, para os seguintes problemas de PL:

   (a) um problema inviável;

   (b) um problema que gera uma solução básica viável inicial para a fase 2, com alguma variável artificial como variável básica; e

   (c) um problema ilimitado;

4. Implemente, em MATLAB ou em OCTAVE, o algoritmo simplex revisado fases 1 e 2, com a regra de Bland, para o problema da dieta exemplificado no Capítulo 3.

5. Dada uma excelente implementação do Algoritmo 13.1.2, sob condições razoáveis de trabalho (energia, digitação dos dados), este algoritmo resolve qualquer PPL no formato padrão. Verdadeiro ou falso? Justifique.

Capítulo 14

Método Afim-Escala: algoritmo mestre

Continuamos com a etapa de estudarmos alguns métodos para resolver problemas de PL. O objetivo deste capítulo é enunciar um algoritmo mestre na tentativa de exprimir algoritmicamente as idéias do método afim-escala (veja Dikin [15] e [16], Barnes [3] e Vanderbei, Meketon e Freedman [58]). No próximo capítulo, enunciaremos o algoritmo afim-escala e demonstraremos sua convergência.

14.1 Algoritmo mestre

Consideremos o PPL primal no formato padrão

$$\begin{array}{lll} (P) & \textit{minimizar} & c^T x \\ & \textit{sujeito a:} & Ax = b \\ & & \ \ x \geq 0, \end{array}$$

onde são dados uma matriz $A \in R^{m \times n}$ e vetores $b \in R^m$ e $c \in R^n$, com $0 < m < n$.

Sem perda de generalidade, conforme Capítulo 2, consideremos a matriz A de posto completo. Além disso, sem perda de generalidade, conforme Proposição 10.2.3, supomos que o vetor custo c não está no espaço linha da matriz tecnológica A, uma vez que o método afim-escala trata com pontos interiores viáveis.

Atenção: Que tal fazer o exercício 1 agora?

As hipóteses para este capítulo são:

(H_1) O conjunto viável $\mathcal{X}$ é limitado.

(H_2) O conjunto de pontos interiores viáveis $\mathcal{X}^0$ é não vazio.

(H_3) É dado um ponto interior viável inicial $x^0 \in \mathcal{X}^0$.

A hipótese (H_1) é muito forte e não é necessária (vide [59]), mas ela simplifica o processo construtivo que pretendemos desenvolver para o método afim-escala. As hipóteses (H_2) e (H_3) são usuais para métodos que geram uma seqüência de pontos interiores viáveis.

Pelas hipóteses (H_1) e (H_2), podemos usar o Teorema 10.3.1 e concluir que o conjunto de soluções ótimas $\mathcal{X}(\mathcal{P})$ é não vazio e limitado. Observe, também, que o vetor do lado direito b é não nulo, porque se $\mathcal{X}^0$ é não vazio, então o conjunto viável $\mathcal{X}$, para $b = 0$, é um cone convexo, logo, não é um conjunto limitado.

Atenção: Que tal fazer o exercício 2 agora?

Considere o problema (P) e suponha (H_1) e (H_2). A idéia do método afim-escala baseia-se na idéia de Dikin [15]. Isto é, o método de Dikin consiste em caminhar pelo interior de um conjunto poliedral de um PPL (P), através de pontos interiores viáveis gerados pela solução de uma seqüência de subproblemas (P_k), $k = 0, 1, \ldots$, a saber:

$$\begin{array}{lll} (P_k) & \textit{minimizar} & c^T x \\ & \textit{sujeito a:} & Ax = b \\ & & (x - x^k)^T X_k^{-2} (x - x^k) \leq 1, \end{array}$$

onde $x^k \in \mathcal{X}^0$ é o centro do maior elipsóide simples e $X_k = diag(x^k)$.

A seguir vamos enunciar uma proposição que afirma podermos desconsiderar a restrição $x \geq 0$ do problema (P_k). A demonstração pode ser encontrada na página 175 em [3].

Proposição 14.1.1 *Considere $x^k \in \mathcal{X}^0$ e $X_k = diag(x^k)$. Então, o elipsóide*

$$\{x \in R^n;\ (x - x^k)^T X_k^{-2} (x - x^k) \leq 1\}$$

está contido no ortante não negativo em R^n.

Demonstração: Suponha por absurdo que $x_{j_0} < 0$ para algum j_0, com $j_0 = 1, 2, \ldots, n$. Então, como

$$\sum_{j=1}^{n} \frac{(x_j - x_j^k)^2}{(x_j^k)^2} = (x - x^k)^T X_k^{-2}(x - x^k) \leq 1,$$

segue-se que

$$1 \geq \sum_{j=1}^{n} \frac{(x_j - x_j^k)^2}{(x_j^k)^2} \geq \frac{(x_{j_0} - x_{j_0}^k)^2}{(x_{j_0}^k)^2} > 1,$$

o que é um absurdo. Isto finaliza a demonstração. ∎

Assim, Dikin propõe resolver um PPL por uma seqüência de problemas de Programação Não Linear.

Atenção: Que tal fazer o exercício 3 agora?

A seguir, enunciamos um algoritmo mestre em uma tentativa de exprimir algoritmicamente as idéias do método de Dikin.

Algoritmo 14.1.2 *Mestre.*

Dado: x^0 é um ponto interior viável inicial, isto é, $x^0 \in \mathcal{X}^0$.
$k := 0$.
REPITA
  Obtenha $x^{k+1} \in \mathcal{X}$ resolvendo o subproblema (P_k).
  $k := k + 1$.
ATÉ QUE 'convirja'.

Atenção: Que tal fazer o exercício 4 agora?

No restante deste capítulo, construiremos o algoritmo afim-escala baseado neste algoritmo mestre. A hipótese (H_3) afirma que $x^0 \in \mathcal{X}^0$ é dado. Assim, a nossa pergunta é a seguinte:

como resolver o subproblema (P_k), $k = 0, 1, \ldots$?

14.2 Resolução de (P_k)

Consideremos o problema de PL (P). Para o método afim-escala, resolvemos o problema (P) através de uma seqüência de subproblemas (P_k), $k = 0, 1, \ldots$.

Com o intuito de responder a nossa pergunta de como resolver o subproblema (P_k), $k = 0, 1, \ldots$, enunciamos o seguinte resultado, cuja demonstração pode ser encontrada na página 35 em [23].

Proposição 14.2.1 *Suponha que $c \in R^n$ não está no espaço linha de A. Então, o problema*

$$\begin{array}{ll} \textit{minimizar} & c^T h \\ \textit{sujeito a:} & Ah = 0 \\ & \|h\| \leq 1, \end{array}$$

possui a solução,

$$\hat{h} = -\frac{c_p}{\|c_p\|},$$

onde c_p é o vetor c projetado no espaço nulo da matriz A.

Demonstração: Por hipótese, $\|c_p\| > 0$. Pela Proposição 7.1.2, $c = c_p + c_{\tilde{p}}$, onde $c_p \in \mathcal{N}(A)$ e $c_{\tilde{p}} \in \mathcal{R}(A^T)$ é ortogonal ao espaço nulo de A. Para qualquer vetor unitário h em $\mathcal{N}(A)$, $c^T h = c_p^T h$ porque $c_{\tilde{p}}^T h = 0$. Daí, usando a desigualdade de Cauchy-Schwarz e $\|h\| = 1$,

$$|c^T h| = |c_p^T h| \leq \|c_p\| \|h\| = \|c_p\|,$$

que pela definição de valor absoluto, é equivalente a

$$-\|c_p\| \leq c^T h \leq \|c_p\|.$$

Logo, para o problema de minimização, a solução é:

$$\hat{h} = -\frac{c_p}{\|c_p\|},$$

finalizando a demonstração. ∎

Atenção: Que tal fazer o exercício 5 agora?

Responderemos a nossa pergunta de como resolver o subproblema (P_k), $k = 0, 1, \ldots$, com a próxima proposição, cuja demonstração pode ser encontrada nas páginas 19 a 22 em [39].

Proposição 14.2.2 *Suponha que $c \in R^n$ não está no espaço linha de A. Então, o problema (P_k), $k = 0, 1, \ldots$, possui a solução*

$$\hat{x} = X_k(e - \frac{P_{\bar{A}}\bar{c}}{\|P_{\bar{A}}\bar{c}\|}),$$

onde $x^k \in \mathcal{X}^0$, $X_k = diag(x^k)$, $\bar{c} = X_k c$, $\bar{A} = AX_k$ e $P_{\bar{A}}$ é a matriz de projeção no espaço nulo de $\bar{A}$.

Demonstração: Considere o problema (P_k), $k = 0, 1, \ldots$. Fazendo uma mudança de escala

$$x = X_k \bar{x},$$

onde $X_k = diag(x^k)$ e $\bar{x} \in R^n$, transformamos o elipsóide com centro em x^k em uma bola com centro no ponto $e = [1, 1, \ldots, 1]^T \in R^n$, porque o ponto $x^k = X_k e$. Ou seja, substituindo $x = X_k \bar{x}$ no problema (P_k), obtemos o seguinte problema de Programação Não Linear:

$$\begin{array}{lll} (\bar{P}_k) & minimizar & \bar{c}^T \bar{x} \\ & sujeito\ a: & \bar{A}\bar{x} = b \\ & & \|\bar{x} - e\| \leq 1, \end{array}$$

onde $\bar{c} = X_k c$, $\bar{A} = AX_k$ e $\bar{x} \in R^n$, pois

$$c^T x = c^T(X_k \bar{x}) = (X_k c)^T \bar{x} = \bar{c}^T \bar{x},$$

$$b = Ax = A(X_k \bar{x}) = (AX_k)\bar{x} = \bar{A}\bar{x},$$

$$\begin{gathered} (x - x^k)^T X_k^{-2} (x - x^k) \leq 1 \\ (X_k \bar{x} - X_k e)^T X_k^{-2} (X_k \bar{x} - X_k e) \leq 1 \\ (X_k(\bar{x} - e))^T X_k^{-2} (X_k(\bar{x} - e)) \leq 1 \\ (\bar{x} - e)^T X_k X_k^{-2} X_k (\bar{x} - e) \leq 1 \end{gathered}$$

$$(\bar{x}-e)^T(\bar{x}-e) = \|\bar{x}-e\|^2 \leq 1 \Rightarrow \|\bar{x}-e\| \leq 1.$$

Temos que o problema $(\bar{P}_k)$ é o problema (P_k) com uma mudança de escala definida por X_k^{-1}, que minimiza o valor de uma função linear na interseção do hiperplano $\{\bar{x} \in R^n;\ \bar{A}\bar{x} = b\}$ e da bola centrada no vetor de uns $e \in R^n$. Segue-se que o vetor e é um ponto interior viável para o subproblema $(\bar{P}_k)$, $k = 0, 1, \ldots$, pois

$$\bar{A}e = (AX_k)e = A(X_k e) = Ax^k = b \ \text{ e } \ \|e - e\| = 0 \leq 1.$$

Tomando uma direção qualquer $h = \bar{x} - e$, conforme Proposição 8.2.3(a), obtemos

$$\bar{c}^T\bar{x} = \bar{c}^T(e+h) = \bar{c}^Te + \bar{c}^Th,$$

$$b = \bar{A}\bar{x} = \bar{A}(e+h) = \bar{A}e + \bar{A}h = b + \bar{A}h \Rightarrow \bar{A}h = 0,$$

$$\|\bar{x} - e\| \leq 1 \Rightarrow \|h\| \leq 1.$$

Daí, uma vez que $\bar{c}^Te$ é uma constante, obtemos o problema

$$\begin{array}{ll} \textit{minimizar} & \bar{c}^Th \\ \textit{sujeito a:} & \bar{A}h = 0 \\ & \|h\| \leq 1, \end{array}$$

cuja solução é

$$\hat{h} = -\frac{P_{\bar{A}}\bar{c}}{\|P_{\bar{A}}\bar{c}\|},$$

conforme Proposição 14.2.1.

Então,

$$\bar{x} = e + \hat{h} = e - \frac{P_{\bar{A}}\bar{c}}{\|P_{\bar{A}}\bar{c}\|},$$

é a solução para $(\bar{P}_k)$. Finalmente, a solução para o problema (P_k) é obtida reescalando a solução em $(\bar{P}_k)$,

$$\hat{x} = X_k\bar{x} = X_k(e - \frac{P_{\bar{A}}\bar{c}}{\|P_{\bar{A}}\bar{c}\|}).$$

Isto finaliza a demonstração. ∎

Esta proposição esclarece os termos 'afim' e 'escala'. Ou seja, a idéia de Dikin para resolver o problema original (P), é resolver uma seqüência de subproblemas (P_k), $k = 0, 1, \ldots$. Resolver (P_k) significa fazer uma mudança de escala para transformar elipsóide em bola através de uma transformação afim.

Conforme Proposição 14.2.1 e Proposição 8.2.3(a), minimizar uma função linear em uma bola unitária interseção com restrições de igualdade é tomar uma direção no espaço afim (neste caso, o termo afim é posterior aos trabalhos de Dikin, designado para métodos de pontos interiores tipo Gonzaga - veja Todd em [56]) definido por $Ax = b$, através de uma direção de descida no espaço nulo de A.

Atenção: Que tal fazer o exercício 6 agora?

Uma vez que o subproblema (P_k), $k = 0, 1, \ldots$, está definido para o maior elipsóide simples, isto é, tangente aos eixos coordenados no primeiro ortante, devemos descartar a possibilidade de obtermos uma solução neste subproblema com alguma coordenada nula, tal que esta solução não seja uma solução ótima para o problema (P). Descartaremos esta possibilidade no próximo teorema, cuja demonstração pode ser encontrada nas páginas 116 e 117 em [52].

Teorema 14.2.3 *Considere x^* uma solução ótima para o subproblema (P_k), qualquer que seja k, $k = 0, 1, \ldots$. Se $x^*_j = 0$ para algum j, $j = 1, 2, \ldots, n$, então x^* resolve o problema (P).*

Demonstração: Fixe arbitrariamente k, $k = 0, 1, \ldots$, e considere o subproblema (P_k). Suponha x^* uma solução ótima para (P_k) tal que $x^*_j = 0$ para algum j, $j = 1, 2, \ldots, n$. Pela definição de matriz de projeção e das mudanças de escala $\bar{c} = X_k c$ e $\bar{A} = AX_k$, para $X_k = diag(x^k)$ com $x^k \in \mathcal{X}^0$, temos que

$$P_{\bar{A}}\bar{c} = (I - \bar{A}^T(\bar{A}\bar{A}^T)^{-1}\bar{A})\bar{c}$$

$$P_{\bar{A}}\bar{c} = (I - (AX_k)^T(AX_k(AX_k)^T)^{-1}AX_k)X_k c$$

$$P_{\bar{A}}\bar{c} = (I - X_k A^T(AX_k^2A^T)^{-1}AX_k)X_k c$$

$$P_{\bar{A}}\bar{c} = X_k c - X_k A^T(AX_k^2A^T)^{-1}AX_k^2 c$$

$$P_{\bar{A}}\bar{c} = X_k(c - A^T(AX_k^2A^T)^{-1}AX_k^2c).$$

Tomando

$$y = (AX_k^2A^T)^{-1}AX_k^2c \quad \text{e} \quad s = c - A^Ty,$$

obtemos

$$P_{\bar{A}}\bar{c} = X_k s.$$

Então, usando a Proposição 14.3.2,

$$x^* = X_k(e - \frac{P_{\bar{A}}\bar{c}}{\|P_{\bar{A}}\bar{c}\|}) = X_k(e - \frac{X_k s}{\|X_k s\|}) = X_k e - \frac{X_k^2 s}{\|X_k s\|}.$$

Portanto,

$$x^* = x^k - \frac{X_k^2 s}{\|X_k s\|}.$$

Assim, usando a hipótese,

$$0 = x_j^* = x_j^k - \frac{(x_j^k)^2 s_j}{\|X_k s\|}.$$

Então, $x_j^k s_j = \|X_k s\|$. Daí, $x_i^k s_i = 0$ para todo $i \neq j$. Como $x_i^k > 0$ pela definição de X_k, obtemos $s_i = 0$ para todo $i \neq j$. Logo, $s \geq 0$. Tomando y uma solução dual com s uma folga dual viável, isto é

$$s = c - A^Ty \quad \text{e} \quad s \geq 0$$

e, usando a condição de folga complementar $(x^*)^T s = 0$, segue-se pelo Teorema 10.1.4 que x^* é uma solução ótima para o problema primal (P). Isto finaliza a demonstração. ∎

Agora estamos prontos para enunciar alguns algoritmos da família afim-escala. É o que faremos no próximo capítulo.

Atenção: Que tal fazer o exercício 7 agora?

14.3 Exercícios

1. Considere o PPL no formato padrão

$$\begin{array}{ll} \textit{minimizar} & x_1 + x_2 + x_3 \\ \textit{sujeito a}: & x_1 + x_2 + x_3 \quad = 3 \\ & x_1,\ x_2,\ x_3 \geq 0. \end{array}$$

   Encontre o conjunto viável e o conjunto de soluções ótimas. Verifique que o vetor custo está no espaço linha da matriz tecnológica.

2. Desenhe o conjunto viável para cada um dos problemas de PL, a saber:

   (a)
$$\begin{array}{ll} \textit{min} & x_1 \\ \textit{s. a}: & x_1 - x_2 \quad = 2 \\ & x_1,\ x_2 \geq 0. \end{array}$$

   (b)
$$\begin{array}{lrl} \textit{min} & x_1 - x_3 & \\ \textit{s. a}: & x_1 + x_2 + x_3 & = 3 \\ & x_3 & = 3 \\ & x_1,\ x_2,\ x_3 \geq 0. & \end{array}$$

   (c)
$$\begin{array}{ll} \textit{min} & x_1 \\ \textit{s. a}: & x_1 - x_2 \quad = 0 \\ & x_1,\ x_2 \geq 0. \end{array}$$

   (d)
$$\begin{array}{lrl} \textit{min} & x_1 + x_2 + x_3 & \\ \textit{s. a}: & x_1 + x_2 + x_3 & = 3 \\ & x_3 & = 1 \\ & x_1,\ x_2,\ x_3 \geq 0. & \end{array}$$

   Por que não trataremos estes problemas para suas resoluções através do método afim-escala? Observe que no item (c), justificar sua resposta pelo fato de que o vetor do lado direito é nulo está correto, mas lembre-se de que isto é uma conseqüência das hipóteses (H_1) e (H_2). Já no item (d), observe que a solução do problema é trivial se supormos um ponto interior viável conhecido.

3. Desenhe a região viável para o problema de Programação Não Linear

$$\begin{array}{lcr} min & x_1 - 2x_2 & \\ s.\ a: & x_1 + x_2 & = 2 \\ & \frac{(x_1-1,5)^2}{(1,5)^2} + \frac{(x_2-0,5)^2}{(0,5)^2} \leq 1. & \end{array}$$

4. Resolva graficamente o PPL

$$\begin{array}{lcr} min & x_1 & \\ s.\ a: & x_1 + x_2 & = 2 \\ & x_1,\ x_2 \geq 0, & \end{array}$$

pelo algoritmo mestre, tomando $x^0 = [1,5,\ 0,5]^T \in R^2$ e um critério de parada arbitrário.

5. Resolva o problema

$$\begin{array}{lcr} min & x_1 - 2x_2 & \\ s.\ a: & x_1 + x_2 & = 0 \\ & x_1^2 + x_2^2 \leq 1. & \end{array}$$

6. Resolva o problema

$$\begin{array}{lcr} min & x_1 - 2x_2 & \\ s.\ a: & x_1 + x_2 & = 2 \\ & \frac{(x_1-1,5)^2}{(1,5)^2} + \frac{(x_2-0,5)^2}{(0,5)^2} \leq 1. & \end{array}$$

Aqui $x^k = [1,5,\ 0,5]^T \in R^2$. Assim, quem é X_k, X_k^{-1} e X_k^{-2}?

7. Resolva o problema

$$\begin{array}{lcr} min & x_2 & \\ s.\ a: & x_1 & = 1 \\ & x_1,\ x_2 \geq 0. & \end{array}$$

através do subproblema (P_k), para $k = 0$, tal que $x^k = x^0 = [1,1]^T$. Qual é a solução ótima para o problema (P_0)? E para o problema original?

Referências Bibliográficas

[1] K. M. Anstreicher. *Linear programming in* $O((n^3/\ln(n))L)$ *operations.* SIAM Journal on Optimization Vol. 9, N. 4, pp. 803-812, 1999.

[2] A. Arbel. *Exploring Interior-Point Linear Programming: Algorithms and Software.* Foundations of Computing Series, MIT Press, 1993.

[3] E. R. Barnes. *A variation on Karmarkar's algorithm for solving Linear Programming problems.* Mathematical Programming 36 (1986) 174-182.

[4] D. Bayer e J. C. Lagarias. *The nonlinear geometry of linear programming, I. Affine and projective scaling trajectories, II. Legendre transform coordinates and central trajectories.* Trans. Amer. Math. Soc. 314 (1989), 499-581.

[5] M. S. Bazaraa, J. J. Jarvis e H. D. Sherali. *Linear Programming and Network Flows.* 2^a edição, John Wiley & Sons, 1990.

[6] M. S. Bazaraa, H. D. Sherali e C. M. Shetty. *Nonlinear Programming: Theory and Algorithms.* 2^a edição, John Wiley & Sons, 1993.

[7] E. M. L. Beale. *Cycling in the dual simplex algorithm.* Naval Research Logistics Quarterly 2, 269-275, 1955.

[8] R. E. Bixby, J. W. Gregory, I. J. Lustig, R. E. Marsten e D. F. Shanno. *Very large-scale linear programming: a case study in combining interior point and simplex methods.* Operations Research Vol. 40, n. 5, 885-897, 1992.

[9] R. G. Bland. *New finite pivoting rules for the simplex method.* Mathematics of Operations Research Vol. 2, n. 2, 103-107, 1977.

[10] J. L. Boldrini, S. I. R. Costa, V. L. Figueiredo e H. G. Wetzler. *Álgebra Linear.* 3^a edição, Harbra, 1984.

[11] P. F. Bregalda, A. A. F. de Oliveira e C. T. Bornstein. *Introdução à Programação Linear.* 3^a edição, Campus, 1988.

[12] A. Cobham. *The Intrinsic Computational Difficulty of Functions.* Em Proc. Int. Cong. Logic Methodology, ed. Bar-Hillel, North Holland, 24-30 (1964).

[13] G. B. Dantzig. *Maximization of a linear function of variables subject to linear inequalities.* Activity Analysis of Production and Allocation, ed. T. C. Koopmans, John Wiley, 339-347, 1951.

[14] G. B. Dantzig. *Linear Programming and Extensions.* Princeton University Press, 10^a ed., 1993.

[15] I. I. Dikin. *Iterative solution of problems of linear and quadratic programming.* Soviet Math. Dokl. 8 (1967) 674-675.

[16] I. I. Dikin. *On the convergence of an iterative process.* Upravlyaemye Sistemy 12 (1974) 54-60. (Em Russo.)

[17] J. Edmonds. *Paths, Trees and Flowers.* Canad. J. Math. 17, 449-467 (1965).

[18] S.-C. Fang e S. Puthenpura. *Linear Optimization and Extensions: theory and algorithms.* Prentice Hall, 1993.

[19] M. C. Goldbarg e H. P. L. Luna. *Otimização Combinatória e Programação Linear: Modelos e Algoritmos.* Campus, 2000.

[20] D. Goldfarb e J. K. Reid. *A practicable steepest-edge simplex algorithm.* Mathematical Programming 12 (1977) 361-371.

[21] D. Goldfarb e M. J. Todd. *Linear Programming.* G. L. Nemhauser et al. editores, Handbooks in OR & MS, Vol. 1, Capítulo II, Elsevier Science Publishers, 1989.

[22] C. C. Gonzaga. *An algorithm for solving linear programming problems in $O(n^3L)$ operations.* Editado por N. Megiddo, Progress in Mathematical Programming: Interior Point and Related Methods, Springer Verlag, Cap. 1, 1989.

[23] C. C. Gonzaga. *Algoritmos de Pontos Interiores para Programação Linear.* 17º Colóquio Brasileiro de Matemática, IMPA/CNPq, Minicurso, 1989.

[24] C. C. Gonzaga. *On lower bound updates in primal potential reduction methods for linear programming.* Mathematical Programming 52 (1991) 415-428.

[25] C. C. Gonzaga. *Path-following methods for linear programming.* SIAM Review Vol. 34, N. 2, 167-224, 1992.

[26] C. C. Gonzaga. *On the Complexity of Linear Programming.* Resenhas IME-USP, 1995, Vol. 2, N. 2, 197-207.

[27] J.-B. Hiriart-Urruty e C. Lemaréchal. *Convex Analysis and Minimization Algorithms I.* Springer-Verlag, 1993.

[28] N. Karmarkar. *A new polynomial time algorithm for linear programming.* Combinatorica 4 (1984) 373-395.

[29] L. G. Khachiyan. *A polynomial algorithm in linear programming.* Soviet Mathematics Doklady, 20:191-194, 1979.

[30] L. G. Khachiyan. *Polynomial algorithms in linear programming.* U.S.S.R. Comput. Maths. Math. and Phys. Vol. 20, N. 1, 53-72, 1980.

[31] V. Klee e G. J. Minty. *How Good is the Simplex Algorithm?.* Em Inequalities III, ed. O. Shisha, Academic Press, 159-175, 1972.

[32] M. Kojima, N. Megiddo e S. Mizuno. *A primal-dual infeasible-interior-point for linear programming.* Mathematical Programming 61 (1993) 263-280.

[33] E. L. Lima. *Curso de Análise.* Volume 2, Projeto Euclides, IMPA/CNPq, 1981.

[34] I. J. Lustig, R. E. Marsten e D. F. Shanno. *Computational experience with a primal-dual interior point method for linear programming.* Linear Algebra and its Applications 152 (1991) 191-222.

[35] H. V. Machado. *Programação Linear.* 10º Colóquio Brasileiro de Matemática, Poços de Caldas/MG, 1975.

[36] N. Maculan e M. V. F. Pereira. *Programação Linear*. Atlas, 1980.

[37] N. Maculan. *Programação Linear: Método do Simplex*. Notas de aula em 1998.

[38] R. E. Marsten, R. Subramanian, M. Saltzman, I. J. Lustig e D. F. Shanno. *Interior point methods for linear programming: Just call Newton, Lagrange, and Fiacco and McCormick!*. Interfaces 20 (1990) 105-116.

[39] C. A. de J. Martinhon. *Programação Linear e Algoritmos de Pontos Interiores: uma introdução*. VI Semana do Instituto de Matemática e Física da UFG (Minicurso), 27 páginas, 1991.

[40] N. Megiddo. *Pathways to the optimal set in linear programming*. Editado por N. Megiddo, Progress in Mathematical Programming: Interior Point and Related Methods, Springer Verlag, Cap. 8, 1989.

[41] M. A. F. Menezes e R. Vieira. *Um protótipo parcial para um sistema de rações para bovinos*. II Encontro de Matemática Aplicada e Computacional em Brasília, 2001.

[42] M. A. F. Menezes. *Um algoritmo de ponto-interior-inviável com complexidade $O(\sqrt{n}L)$ iterações para programação linear*. Tese de Doutorado pela COPPE/UFRJ, defendida sob a orientação do Professor Clóvis C. Gonzaga em 1998.

[43] S. Mizuno. *Polynomiality of infeasible-interior-point algorithms for linear programming*. Mathematical Programming 67 (1994) 109-119.

[44] K. G. Murty. *Linear Programming*. John Wiley & Sons, 1983.

[45] A. A. Namen e C. T. Bornstein. *Uma ferramenta para avaliação de resultados de diversos modelos de otimização de dietas*. Pesquisa Operacional, v. 24, n. 3, p. 445-465, 2004.

[46] G. L. Nemhauser e L. A. Wolsey. *Integer and Combinatorial Optimization*. John Wiley & Sons, 1988.

[47] Y. Nesterov. *An old and new approach to nonlinear programming*. Mathematical Programming 79 (1997) 285-297.

[48] F. Potra. *An infeasible-interior-point predictor-corrector algorithm for linear programming*. SIAM Journal on Optimization Vol. 6, N. 1, pp. 19-32, 1996.

[49] F. Potra. *A quadratically convergent predictor-corrector method for solving linear programs from infeasible starting points*. Mathematical Programming 67 (1994) 383-406.

[50] A. Ravindran, D. T. Phillips e J. J. Solberg. *Operations Research: principles and practice*. 2^a edição, John Wiley and Sons, 1987.

[51] J. Renegar. *A polynomial-time algorithm, based on Newton's method, for linear programming*. Mathematical Programming 40 (1988) 59-93.

[52] R. Saigal. *Linear Programming: A Modern Integrated Analysis*. Kluwer Academic Publishers, 1995.

[53] R. Shamir. *The efficiency of the simplex method: a survey*. Management Science 33, 301-334, 1987.

[54] N. Z. Shor. *Utilization of the operation of space dilation in the minimization of convex functions*. Kibernetika, N. 1, 6-12, 1970 (traduzido em Cybernetics 6 (1970) 7-15).

[55] G. Strang. *Linear Algebra and its Applications*. 3^a edição, Harcourt Brace Jovanovich, 1988.

[56] M. J. Todd. *Potential-reduction methods in mathematical programming*. Mathematical Programming 76 (1996) 3-45.

[57] P. Vaidya. *An algorithm for linear programming which requires $O(((m+n)n^2 + (m+n)^{1.5}n)L)$ arithmetic operations*. Mathematical Programming 47 (1990) 175-201.

[58] R. J. Vanderbei, M. S. Meketon e B. A. Freedman. *A Modification of Karmarkar's Linear Programming Algorithm*. Algorithmica 1 (1986) 395-407.

[59] R. J. Vanderbei e J. C. Lagarias. *I. I. Dikin's Convergence Result for the Affine-Scaling Algorithm*. Contemporary Mathematics, Vol. 114, pp. 109-119, 1990.

[60] J. von Neumann. *A certain Zero-Sum Two-Person Game Equivalent to the Optimal Assignment Problem*. Em Contributions to the Theory of Games II, ed. H. W. Kuhn e A. W. Tucker. Princeton U. Press, 1953.

[61] Y. Ye. *An $O(\sqrt{n}L)$-Iteration Combined Phase I-Phase II Potential Reduction Algorithm for Linear Programming*. Department of Management Sciences, The University of Iowa, Iowa City, Iowa, 1992.

[62] Y. Ye, M. J. Todd e S. Mizuno. *An $O(\sqrt{n}L)$-iteration homogeneous and self-dual linear programming algorithm*. Mathematics of Operations Research, Vol. 19, N. 1, 53-67, 1994.

[63] D. B. Yudin e A. S. Nemirovskii. *Informational complexity and efficient methods for the solution of convex extremal problems*. Ekonomika Matematicheskie Metody 12, N. 2, 357-369, 1976 (traduzido em Matekon 13 (1977) 25-45).

[64] Y. Zhang. *On the convergence of a class of infeasible interior-point methods for the horizontal linear complementarity problem*. SIAM Journal on Optimization Vol. 4, N. 1, pp. 208-227, 1994.