INF05010 - Otimização combinatória
Notas de aula

Alysson M. Costa, Luciana Buriol, Marcus Ritt
{amcosta,buriol,mrpritt}@inf.ufrgs.br

27 de Março de 2009

Universidade Federal do Rio Grande do Sul
Instituto de Informática
Departamento de Informática Teórica

Conteúdo

1 Introdução

1.1 Exemplo

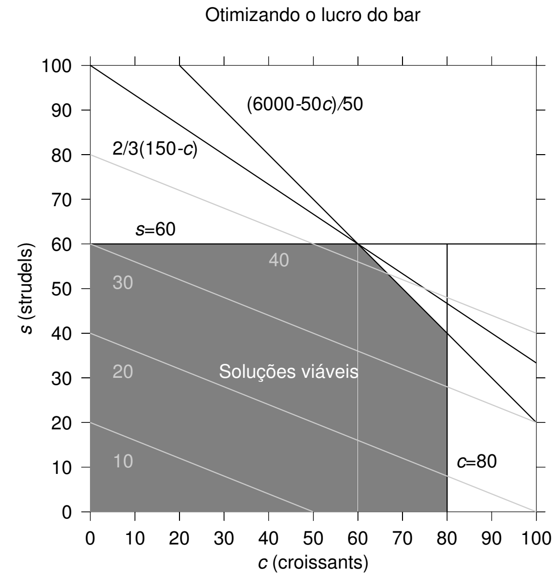
Exemplo 1.1 (No Ildo)
Antes da aula visito o Ildo para tomar um café e comer um Croissant. Ele me conta: "Estou especializado em Croissants e Strudels. Tenho um lucro de 20 centavos por Croissant e 50 centavos por Strudel. Diariamente até 80 clientes compram um Croissant e até 60 um Strudel. Mas infelizmente, o Ildo apenas disponibiliza de 150 ovos e 6 kg de açúcar por dia. Entre outros ingredientes, preciso um ovo e 50g de açúcar para cada Croissant e um ovo e meio e 50g de açúcar para cada Strudel. Agora, professor, quantas Croissants e Strudels devo produzir para obter o maior lucro?"

Sejam c e s o número de Croissants e Strudels, respectivamente. O lucro do Ildo em Reais é $0.2c+0.5s$. Seria ótimo produzir todos 80 Croissants e 60 Strudels, mas uma conta simples mostra que não temos ovos e açúcar suficientes. Para produzir os Croissants e Strudels precisamos $c + 1.5s$ ovos e $50c + 50s$g de açúcar que não podem ultrapassar 150 ovos e 6000g. Com a condição óbvia que $c \geq 0$ e $s \geq 0$ chegamos no seguinte problema de otimização:

$$\begin{aligned}
\textbf{maximiza} \quad & 0.2c + 0.5s \\
\textbf{sujeito a} \quad & c + 1.5s \leq 150 \\
& 50c + 50s \leq 6000 \\
& c \leq 80 \\
& s \leq 60 \\
& c, s \geq 0
\end{aligned}$$

Como resolver esse problema? Com duas variáveis podemos visualizar a situação num grafo com c no eixo x e s no eixo y

No Ildo



que nesse caso permite resolver o problema graficamente. Desenhando diversos *conjuntos de nível* (ingl. *level set*) com valor da função objetivo 10, 20, 30, 40 é fácil observar que o lucro máximo se encontra no ponto $c = s = 60$, e possui um valor de 42 reais.

◊

Isso é um exemplo de um problema de otimização. A forma geral de um *problema de otimização* (ou de *programação matemática*) é

$$\begin{aligned}
\textbf{opt} \quad & f(x) \\
\textbf{sujeito a} \quad & x \in V
\end{aligned}$$

com

- um *objetivo* **opt** $\in \{\max, \min\}$,
- uma *função objetivo* (ou função critério) $f : V \to \mathbb{R}$,
- um conjunto de *soluções viáveis* (ou soluções candidatas) V.

Falamos de um *problema de otimização combinatória*, se V é discreto.
Nessa generalidade um problema de otimização é difícil de resolver. O exemplo 1.1 é um problema de *otimização linear* (ou *programação linear*):

- as variáveis da solução são $x_1, \dots, x_n \in \mathbb{R}$
- a função de otimização é linear em $x_1, \dots, x_n$:

$$f(x_1, \dots, x_n) = c_1x_1 + \cdots + c_nx_n \tag{1.1}$$

- as soluções viáveis são dadas implicitamente por m *restrições lineares*

$$a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \bowtie_1 b_1 \tag{1.2}$$
$$a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \bowtie_2 b_2 \tag{1.3}$$
$$\cdots \tag{1.4}$$
$$a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \bowtie_m b_m \tag{1.5}$$

  com $\bowtie_i \in \{\le, =, \ge\}$.

Exemplo 1.2 (O problema da dieta)
Suponha que temos uma tabela de nutrientes de diferentes tipos de alimentos. Sabendo o valor diário de referência (VDR) de cada nutriente (quantidade de nutriente que deve ser ingerido) e o preço de cada unidade de alimento, qual a dieta ótima, i.e. que contém ao menos o valor diário de referência, mas de menor custo?
Com m nutrientes e n alimentos, seja a_{ij} a quantidade do nutriente i no alimento j, r_i o valor diário de referência do nutriente i e c_j o preço do alimento j. Queremos saber as quantidades x_j de cada alimento que

$$\begin{array}{rl} \textbf{minimiza} & c_1x_1 + \cdots + c_nx_n \\ \textbf{sujeito a} & a_{11}x_1 + \cdots + a_{1n}x_n \ge r_1 \\ & \cdots \\ & a_{m1}x_1 + \cdots + a_{mn}x_n \ge r_m \\ & x_1, \dots, x_n \ge 0 \end{array}$$

$\Diamond$

Exemplo 1.3 (Problema de transporte)
Uma empresa agrária tem m depósitos, cada um com um estoque de a_i ($1 \le i \le m$) toneladas de milho. Ela quer encaminhar b_j ($1 \le j \le n$) toneladas de milho para n clientes diferentes. O transporte de uma tonelada do depósito i para cliente j custa R\$ c_{ij}. Qual seria o esquema de transporte de menor custo?
Como problema de otimização linear, podemos introduzir como variáveis x_{ij} o peso dos produtos encaminhados pelo depósito i para cliente j, e queremos resolver

$$\begin{array}{rll} \textbf{minimiza} & \sum_{ij} c_{ij}x_{ij} & \\ \textbf{sujeito a} & \sum_j x_{ij} \le a_i & \text{para todo fornecedor } i \\ & \sum_i x_{ij} = b_j & \text{para todo cliente } j \\ & x_{ij} \ge 0 & \end{array}$$

Concretamente, suponha que temos a situação da figura 1.1. A figura mostra

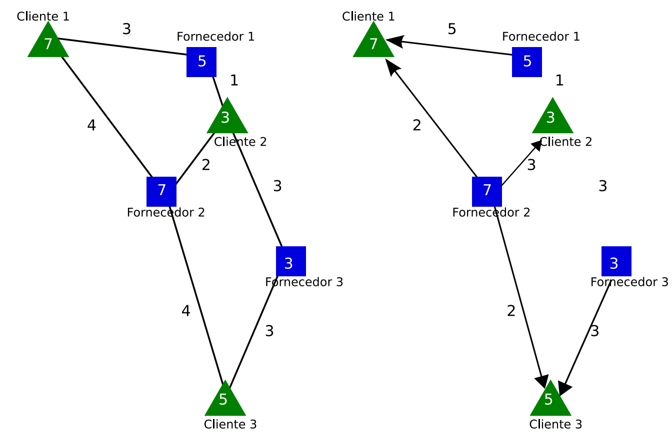


Figura 1.1: Esquerda: Instância do problema de transporte. Direita: Solução ótima dessa instância.

as toneladas disponíveis de cada fornecedor, a demanda (em toneladas) de cada cliente e as distâncias (em km) entre eles. O transporte custa R\$ 1000 por km e tonelada. Observe que um transporte do fornecedor 1 para cliente 3 e fornecedor 3 para cliente 1 não é possível. Nós usaremos uma distância grande de 100km nesses casos (outra possibilidade seria usar restrições $x_{13} = x_{31} = 0$.

$$\begin{array}{rl} \textbf{minimiza} & 3x_{11} + x_{12} + 100x_{13} + 4x_{21} + 2x_{22} \\ & + 4x_{23} + 100x_{31} + 3x_{32} + 3x_{33} \\ \textbf{sujeito a} & x_{11} + x_{12} + x_{13} \le 5 \\ & x_{21} + x_{22} + x_{23} \le 7 \\ & x_{31} + x_{32} + x_{33} \le 3 \\ & x_{11} + x_{21} + x_{31} = 7 \\ & x_{12} + x_{22} + x_{32} = 3 \\ & x_{13} + x_{23} + x_{33} = 5 \\ & x_{ij} \ge 0 \end{array}$$

Qual seria a solução ótima? A figura da direita mostra o número ótimo de toneladas transportadas. O custo mínimo é 46 (em R$ 1000). ◊

Para simplificar a descrição, podemos usar matrizes e vetores. Como exemplo, considere o problema de otimização generalizado visto anteriormente (equações 1.1-1.5). Com $A := (a_{ij}) \in \mathbb{R}^{m\times n}$, $b := (b_i) \in \mathbb{R}^m$, $c := (c_i) \in \mathbb{R}^n$ e $x = (x_i) \in \mathbb{R}^n$ o sistema acima é equivalente a

$$\begin{array}{rll} \textbf{opt} & c^t x & \\ \textbf{sujeito a} & a_i x \bowtie_i b_i & 1 \le i \le m \end{array}$$

(Denotamos com a_i a i-ésima linha e como a^j a j-ésima coluna de A.)

1.2 Formas normais

Conversões

É possível converter

- um problema de minimização para um problema de maximização

$$\min c^t x \iff \max -c^t x$$

- uma restrição $\ge$ para uma restrição $\le$

$$a_i x \ge b_i \iff -a_i x \le -b_i$$

- uma igualdade para desigualdades

$$a_i x = b_i \iff a_i x \le b_i \wedge a_i x \ge b_i$$

Conversões

- uma desigualdade para uma igualdade

$$\begin{aligned} a_i x \le b &\iff a_i x + x_{n+1} = b_i \wedge x_{n+1} \ge 0 \\ a_i x \ge b &\iff a_i x - x_{n+1} = b_i \wedge x_{n+1} \ge 0 \end{aligned}$$

  usando uma nova *variável de folga ou excesso* x_{n+1} (inglês: slack and surplus variables).

- uma variável x_i sem restrições para duas positivas

$$x_i^+ \ge 0 \wedge x_i^- \ge 0$$

  substituindo x_i por $x_i^+ - x_i^-$.

Essas transformações permitem descrever cada problema linear em uma *forma padrão*.

Forma padrão

$$\begin{array}{rl} \textbf{maximiza} & c^t x \\ \textbf{sujeito a} & Ax \le b \\ & x \ge 0 \end{array}$$

As restrições $x \ge 0$ se chamam *triviais*.

Exemplo 1.4

Dado o problema

$$\begin{array}{rl} \textbf{minimiza} & 3x_1 - 5x_2 + x_3 \\ \textbf{sujeito a} & x_1 - x_2 - x_3 \ge 0 \\ & 5x_1 + 3x_2 + x_3 \le 200 \\ & 2x_1 + 8x_2 + 2x_3 \le 500 \\ & x_1, x_2 \ge 0 \end{array}$$

vamos substituir **minimiza** para **maximiza**, converter a primeira desigualdade de $\ge$ para $\le$ e introduzir $x_3 = x_3^+ - x_3^-$ com duas variáveis positivas x_3^+

e x_3^- para obter a forma padrão

$$\begin{aligned} \textbf{maximiza} \quad & -3x_1 + 5x_2 - x_3^+ + x_3^- \\ \textbf{sujeito a} \quad & -x_1 + x_2 + x_3^+ - x_3^- \le 0 \\ & 5x_1 + 3x_2 + x_3^+ - x_3^- \le 200 \\ & 2x_1 + 8x_2 + 2x_3^+ - 2x_3^- \le 500 \\ & x_1, x_2, x_3^+, x_3^- \ge 0 \end{aligned}$$

Em notação matricial temos

$$c = \begin{pmatrix} -3 \\ 5 \\ -1 \\ 1 \end{pmatrix}; b = \begin{pmatrix} 0 \\ 200 \\ 500 \end{pmatrix}; A = \begin{pmatrix} -1 & 1 & 1 & -1 \\ 5 & 3 & 1 & -1 \\ 2 & 8 & 2 & -2 \end{pmatrix}.$$

◊

1.3 Notas históricas

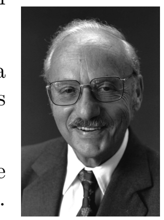
História da programação linear

- Jean Baptiste Joseph Fourier (1826): Método de resolver um sistema de desigualdades (eliminação de Fourier-Motzkin) [5].
- Leonid Kantorovich (1939): Programação linear.
- George Bernard Dantzig (1948): Método Simplex.
- John von Neumann: Dualidade.
- Leonid Khachiyan (1979): Método de ellipsoides.
- Narendra Karmarkar (1984): Métodos de pontos interiores.

Jean Baptiste Joseph Fourier (*1768, +1830)

Pesquisa operacional, otimização e "programação"

- "The discipline of applying advanced analytical methods to help make better decisions" (INFORMS)
- A noção foi criada no segunda guerra mundial, para métodos científicos de análise e predição de problemas logísticos.
- Hoje se aplica para técnicas que ajudam decisões de execução e coordenação de operações em organizações.
- Os problemas da pesquisa operacional são problemas de otimização.
- "Programação" ≠ "Programação"
  - Não se refere à computação: a noção significa "planejamento" ou "agendamento".



George Bernard Dantzig (*1914, +2005)

Técnicas da pesquisa operacional

- Em geral: Técnicas algorítmicas conhecidas como
  - Modelagem matemática (equações, igualdades, desigualdades, modelos probabilísticos,...)
  - Algoritmos gulosos, randômicos, ...; programação dinâmica, linear, convexo, ...
  - Heurísticas e algoritmos de aproximação.
- Algumas dessas técnicas se aplicam para muitos problemas e por isso são mais comuns:
  - Exemplo: Programação linear.

2 O método Simplex

Graficamente, é difícil resolver sistemas com mais de três variáveis. Portanto é necessário achar métodos que permitam resolver sistemas grandes. Um método importante se chama Simplex. Nós vamos estudar esse método primeiramente através da aplicação a um exemplo.

2.1 Um exemplo

Começamos com o seguinte sistema em forma padrão:

Exemplo: Simplex

$$\begin{aligned} \textbf{maximiza} \quad & z = 6x_1 + 8x_2 + 5x_3 + 9x_4 \\ \textbf{sujeito a} \quad & 2x_1 + x_2 + x_3 + 3x_4 \leq 5 \\ & x_1 + 3x_2 + x_3 + 2x_4 \leq 3 \\ & x_1, x_2, x_3, x_4 \geq 0 \end{aligned}$$

Introduzimos variáveis de folga e reescrevemos as equações:

Exemplo: Com variáveis de folga

$$\begin{aligned} \textbf{maximiza} \quad & z = 6x_1 + 8x_2 + 5x_3 + 9x_4 && (2.1) \\ \textbf{sujeito a} \quad & w_1 = 5 - 2x_1 - x_2 - x_3 - 3x_4 && (2.2) \\ & w_2 = 3 - x_1 - 3x_2 - x_3 - 2x_4 && (2.3) \\ & x_1, x_2, x_3, x_4, w_1, w_2 \geq 0 \end{aligned}$$

Nesse exemplo é fácil obter uma solução viável, escolhendo $x_1 = x_2 = x_3 = x_4 = 0$. Podemos verificar que $w_1 = 5$ e $w_2 = 3$ e todas as restrições são respeitadas. O valor da função objetivo seria 0. Uma outra solução viável é $x_1 = 1$, $x_2 = x_3 = x_4 = 0$, $w_1 = 3$, $w_2 = 2$ com valor $z = 6$.
Com 6 variáveis e duas equações independentes o espaço de soluções do sistema de equações lineares dado pelas restrições tem $6 - 2 = 4$ graus de liberdade.

Uma solução viável com no mínimo esse número de *variáveis nulas* (igual a 0) se chama uma *solução básica viável*. Logo nossa primeira solução acima é uma solução básica viável.
A idéia do método Simplex é percorrer soluções básicas viáveis, aumentando em cada passo o valor z da função objetivo.
Logo nosso próximo objetivo é aumentar o valor da função objetivo z. Para esse fim, podemos aumentar o valor das variáveis x_1, x_2, x_3 ou x_4, pois o coeficiente delas é positivo. Escolhemos x_4, porque essa variável tem o maior coeficiente. Não podemos aumentar x_4 arbitrariamente: Para respeitar as restrições $w_1, w_2 \geq 0$ temos os limites

Limites

$$\begin{aligned} w_1 = 5 - 3x_4 \geq 0 \iff x_4 \leq 5/3 \\ w_2 = 3 - 2x_4 \geq 0 \iff x_4 \leq 3/2 \end{aligned}$$

ou seja $x_4 \leq 3/2$. Aumentando x_4 o máximo possível, obtemos $x_4 = 3/2$ e $w_2 = 0$. Os valores das demais variáveis não mudam. Essa solução respeita novamente todas as restrições, e portanto é *viável*. Ainda, como trocamos uma variável nula (x_4) com uma outra não-nula (w_2) temos uma nova solução básica viável

Solução básica viável

$$x_1 = x_2 = x_3 = 0; x_4 = 3/2; w_1 = 1/2; w_2 = 0$$

com valor da função objetivo $z = 13.5$.
O que facilitou esse primeiro passo foi a forma especial do sistema de equações. Escolhemos quatro variáveis independentes (x_1, x_2, x_3 e x_4) e duas variáveis dependentes (w_1 e w_2). Essas variáveis são chamadas *não-básicas* e *básicas*, respectivamente. Na nossa solução básica viável todas variáveis não-básicas são nulas. Logo, pode-se aumentar uma variável não-básica cujo coeficiente na função objetivo seja positivo (para aumentar o valor da função objetivo). Inicialmente tem-se as seguintes variáveis básicas e não-básicas

$$\mathcal{B} = \{w_1, w_2\}; \qquad \mathcal{N} = \{x_1, x_2, x_3, x_4\}.$$

Depois de aumentar x_4 (e consequentemente zerar w_2) podemos escolher

$$\mathcal{B} = \{w_1, x_4\}; \qquad \mathcal{N} = \{x_1, x_2, x_3, w_2\}.$$

A variável x_4 se chama *variável entrante*, porque ela entra no conjunto de variáveis básicas B. Analogamente w_2 se chama *variável sainte*.
Para continuar, podemos reescrever o sistema atual com essas novas variáveis básicas e não-básicas. A segunda restrição 2.3 é fácil de reescrever

$$w_2 = 3 - x_1 - 3x_2 - x_3 - 2x_4 \iff x_4 = 3/2 - 1/2x_1 - 3/2x_2 - 1/2x_3 - 1/2w_2$$

Além disso, temos que reescrever a primeira restrição 2.2, porque a variável básica w_1 depende de x_4 que agora é básica também. Nosso objetivo é escrever todas variáveis básicas em termos de variáveis não-básicas. Para esse fim, podemos usar combinações lineares da linhas, que eliminam as variáveis não-básicas. Em nosso exemplo, a combinação (2.2)$-3/2$(2.3) elimina x_4 e resulta em

$$w_1 - 3/2w_2 = 1/2 - 1/2x_1 + 7/2x_2 + 1/2x_3$$

e colocando a variável não-básica w_2 no lado direto obtemos

$$w_1 = 1/2 - 1/2x_1 + 7/2x_2 + 1/2x_3 + 3/2w_2.$$

Temos que aplicar uma operação semelhante à função objetivo que ainda depende da variável básica x_4. Escolhemos (2.1)$-9/2$(2.3) para obter

$$z = 27/2 + 3/2x_1 - 11/2x_2 + 1/2x_3 - 9/2w_2.$$

Novo sistema

$$\begin{array}{rl}
\textbf{maximiza} & z = 27/2 + 3/2x_1 - 11/2x_2 + 1/2x_3 - 9/2w_2 \\
\textbf{sujeito a} & w_1 = 1/2 - 1/2x_1 + 7/2x_2 + 1/2x_3 + 3/2w_2 \\
& x_4 = 3/2 - 1/2x_1 - 3/2x_2 - 1/2x_3 - 1/2w_2 \\
& x_1, x_2, x_3, x_4, w_1, w_2 \geq 0
\end{array}$$

que obtemos após uma operação de trocar as variáveis x_4 e w_2. Essa operação se chama um *pivot*. Observe que no novo sistema é fácil recuperar toda informação atual: zerando as variáveis não-básicas obtemos diretamente a solução $x_1 = x_2 = x_3 = w_2 = 0$, $w_1 = 1/2$ e $x_4 = 3/2$ com função objetivo $z = 27/2$.
Antes de continuar "pivotando" introduzimos uma forma mais simples de escrever o sistema

Dicionário

$$\begin{array}{rllllll}
 & & & \downarrow & & & \\
 & z & = 27/2 & +3/2x_1 & -11/2x_2 & +1/2x_3 & -9/2w_2 \\
\hline
\leftarrow & w_1 & = 1/2 & -1/2x_1 & +7/2x_2 & +1/2x_3 & +3/2w_2 \\
 & x_4 & = 3/2 & -1/2x_1 & -3/2x_2 & -1/2x_3 & -1/2w_2
\end{array}$$

que se chama *dicionário* (inglês: dictionary).
No próximo passo podemos aumentar somente x_1 ou x_3 porque somente elas têm coeficientes positivos. Aumentado x_1 temos que respeitar $x_1 \leq 1$ (da primeira restrição) e $x_1 \leq 3$ (da segunda). Logo a primeira restrição é mais forte, x_1 é a variável entrante, w_1 a variável sainte, e depois do pivot obtemos

Segundo passo

$$\begin{array}{rllllll}
 & & & & \downarrow & & \\
 & z & = 15 & -3w_1 & +5x_2 & +2x_3 & \\
\hline
 & x_1 & = 1 & -2w_1 & +7x_2 & +x_3 & +3w_2 \\
\leftarrow & x_4 & = 1 & +w_1 & -5x_2 & -x_3 & -2w_2
\end{array}$$

No próximo pivot x_2 entra. A primeira restrição não fornece limite para x_2, porque o coeficiente de x_2 é positivo! Mas a segunda $x_2 \leq 1/5$ e x_4 sai da base. O resultado do pivot é

Terceiro passo

$$\begin{array}{rllllll}
 & & & & & \downarrow & \\
 & z & = 16 & -2w_1 & -x_4 & +x_3 & -2w_2 \\
\hline
 & x_1 & = 12/5 & -3/5w_1 & -7/5x_4 & -2/5x_3 & +1/5w_2 \\
\leftarrow & x_2 & = 1/5 & +1/5w_1 & -1/5x_4 & -1/5x_3 & -2/5w_2
\end{array}$$

O próximo pivot: x_3 entra, x_2 sai:

Quarto passo

$$\begin{array}{llllll}
z & = 17 & -w_1 & -2x_4 & -5x_2 & -4w_2 \\
\hline
x_1 & = 2 & -w_1 & -x_4 & +2x_2 & +w_2 \\
x_3 & = 1 & +w_1 & -x_4 & -5x_2 & -2w_2
\end{array}$$

Agora, todos coeficientes da função objetivo são negativos. Isso significa, que não podemos mais aumentar nenhuma variável não-básica. Como esse sistema é equivalente ao sistema original, qualquer solução tem que ter um valor menor ou igual a 17, pois todas as variáveis são positivas. Logo chegamos no resultado final: a solução

$$w_1 = x_4 = x_2 = w_2 = 0; x_1 = 2; x_3 = 1$$

com valor objetivo 17, é ótima!
Concluímos esse exemplo com mais uma observação. O número de soluções básicas viáveis é limitado. Em nosso exemplo, se escolhemos um subconjunto de quatro variáveis nulas, as duas equações determinam as variáveis restantes. Logo temos no máximo $\binom{6}{4} = 15$ soluções básicas viáveis. Com m equações e n variáveis, uma solução básica viável possui $n-m$ variáveis independentes (que são nulas). Portanto, o número de soluções básicas viáveis é igual a $\binom{n}{n-m}$. Dessa forma, se aumentamos em cada pivot o valor da função objetivo, o método termina em no máximo $\binom{n}{n-m}$ passos.

Exemplo 2.1 (Solução do problema do Ildo)
Exemplo da solução do problema do Ildo na página 9.

$$\begin{array}{rrrr}
 & & & \downarrow \\
z = & 0/1 & +1/5c & +1/2s \\
w_1 = & 150/1 & -1/1c & -3/2s \\
w_2 = & 6000/1 & -50/1c & -50/1s \\
w_3 = & 80/1 & -1/1c & 0/1s \\
\leftarrow w_4 = & 60/1 & 0/1c & -1/1s
\end{array}$$

Pivô s–w_4

$$\begin{array}{rrrr}
 & & \downarrow & \\
z = & 30/1 & +1/5c & -1/2w_4 \\
\leftarrow w_1 = & 60/1 & -1/1c & +3/2w_4 \\
w_2 = & 3000/1 & -50/1c & +50/1w_4 \\
w_3 = & 80/1 & -1/1c & 0/1w_4 \\
s = & 60/1 & 0/1c & -1/1w_4
\end{array}$$

Pivô c–w_1

$$\begin{array}{rrrr}
z = & 42/1 & -1/5w_1 & -1/5w_4 \\
c = & 60/1 & -1/1w_1 & +3/2w_4 \\
w_2 = & 0/1 & +50/1w_1 & -25/1w_4 \\
w_3 = & 20/1 & +1/1w_1 & -3/2w_4 \\
s = & 60/1 & 0/1w_1 & -1/1w_4
\end{array}$$

Lucro total de R$ 42, com os seguintes valores de variáveis: $c = 60$, $s = 60$, $w_1 = 0$, $w_2 = 0$, $w_3 = 20$ e $w_4 = 0$.
Interpretação das variáveis:

- c: Número de Croissants produzidos.
- s: Número de Strudels produzidos.
- w_1: Número de ovos sobrando: 0.
- w_2: Quantidade de açúcar sobrando: 0 g.
- w_3: Croissants não produzidos (abaixo da demanda): 20.
- w_4: Strudels não produzidos: 0.

◊

2.2 O método resumido

Considerando n variáveis e m restrições:

Sistema inicial

$$\begin{array}{rll}
\textbf{maximiza} & z = \sum_{1\le j\le n} c_j x_j & \\
\textbf{sujeito a} & \sum_{1\le j\le n} a_{ij} x_j \le b_i & 1 \le i \le m \\
 & x_j \ge 0 & 1 \le j \le n
\end{array}$$

Preparação

Introduzimos variáveis de folga

$$\sum_{1\le j\le n} a_{ij} x_j + x_{n+i} = b_i \qquad 1 \le i \le m$$

e escrevemos as variáveis de folga como dependentes das variáveis restantes

$$x_{n+i} = b_i - \sum_{1\le j\le n} a_{ij} x_j \qquad 1 \le i \le m$$

Solução básica viável inicial
Se todos $b_i \geq 0$ (o caso contrário vamos tratar na próxima seção), temos uma solução básica inicial

$$\begin{aligned} x_{n+i} &= b_i \qquad 1 \leq i \leq m \\ x_j &= 0 \qquad 1 \leq j \leq n \end{aligned}$$

Índices das variáveis
Depois do primeiro passo, os conjuntos de variáveis básicas e não-básicas mudam. Seja $\mathcal{B}$ o conjunto dos índices das variáveis básicas (dependentes, em geral não-nulas) e $\mathcal{N}$ o conjunto das variáveis independentes (nulas). No começo temos

$$\mathcal{B} = \{n+1, n+2, \ldots, n+m\}; \qquad \mathcal{N} = \{1, 2, \ldots, n\}$$

A forma geral do sistema muda para

$$\begin{aligned} z &= \bar{z} + \sum_{j \in \mathcal{N}} \bar{c}_j x_j \\ x_i &= \bar{b}_i - \sum_{j \in \mathcal{N}} \bar{a}_{ij} x_j \qquad i \in \mathcal{B} \end{aligned}$$

As barras em cima dos coeficientes enfatizam que eles mudam ao longo da aplicação do método.

Escolher variável entrante
Em cada passo do método Simplex, escolhemos uma variável não-básica x_k, com $k \in \mathcal{N}$ para aumentar o valor objetivo z. Isso somente é possível para os índices j tal que $\bar{c}_j > 0$, i.e.

$$\{j \in \mathcal{N} \mid \bar{c}_j > 0\}.$$

Escolhemos um k desse conjunto, e x_k é a variável entrante. Uma heurística simples é a *regra do maior coeficiente*, que escolhe

$$k = \operatorname{argmax}\{\bar{c}_j \mid \bar{c}_j > 0, j \in \mathcal{N}\}$$

Aumentar a variável entrante
Seja x_k a variável entrante. Se aumentamos x_k para um valor positivo, as variáveis básicas têm novos valores

$$x_i = \bar{b}_i - \bar{a}_{ik} x_k \qquad i \in \mathcal{B}.$$

Temos que respeitar $x_i \geq 0$ para $1 \leq i \leq n$. Cada equação com $\bar{a}_{ik} > 0$ fornece uma cota superior para x_k:

$$x_k \leq \bar{b}_i / \bar{a}_{ik}.$$

Logo podemos aumentar x_k ao máximo um valor

$$\alpha := \min_{\substack{i \in \mathcal{B} \\ \bar{a}_{ik} > 0}} \frac{\bar{b}_i}{\bar{a}_{ik}} > 0.$$

Podemos escolher a variável sainte entre os índices

$$\{i \in \mathcal{B} \mid \bar{b}_i / \bar{a}_{ik} = \alpha\}.$$

2.3 Interpretação geométrica

2.4 Sistemas ilimitados

Como pivotar?

- Considere o sistema

$$\begin{array}{llll} z & = 24 & -x_1 & +2x_2 \\ \hline x_3 & = 2 & -x_1 & +x_2 \\ x_4 & = 5 & +x_1 & +4x_2 \end{array}$$

- Qual a próxima solução básica viável?
- A duas equações não restringem o aumento de x_2: existem soluções com valor ilimitado.

2.5 Encontrar uma solução inicial

Solução básica inicial

- Nosso problema inicial é

$$\begin{aligned} \textbf{maximiza} \quad & z = \sum_{1\le j\le n} c_j x_j \\ \textbf{sujeito a} \quad & \sum_{1\le j\le n} a_{ij}x_j \le b_i && 1 \le i \le m \\ & x_i \ge 0 && 1 \le i \le m \end{aligned}$$

- com dicionário inicial

$$\begin{aligned} z &= \sum_{1\le j\le n} c_j x_j \\ x_{n+i} &= b_i - \sum_{1\le j\le n} a_{ij}x_j \qquad 1 \le i \le m \end{aligned}$$

Solução básica inicial

- A solução básica inicial desse dicionário é

$$x = (0 \cdots 0\, b_1 \cdots b_m)^t$$

- O que acontece se existe um $b_i < 0$?
- A solução básica não é mais viável! Sabe-se disso porque pelo menos uma variável básica terá valor negativo.

Sistema auxiliar

- Um método para resolver o problema: resolver outro programa linear
  - cuja solução fornece uma solução básica viável do programa linear original e
  - que tem uma solução básica viável simples, tal que podemos aplicar o método Simplex.

$$\begin{aligned} \textbf{maximiza} \quad & z = -x_0 \\ \textbf{sujeito a} \quad & \sum_{1\le j\le n} a_{ij}x_j - x_0 \le b_i && 0 \le i \le m \\ & x_i \ge 0 && 0 \le i \le n \end{aligned}$$

Resolver o sistema auxiliar

- É fácil achar uma solução viável do sistema auxiliar
- Escolhe $x_i = 0$, para todos $1 \le i \le n$.
- Escolhe x_0 suficientemente grande: $x_0 \ge \max_{1\le i\le m} -b_i$.
- Isso corresponde com um primeiro pivot com variável entrante x_0 após introduzir as variáveis de folga
  - Podemos começar com a solução não-viável $x_0 = x_1 = \ldots = x_n = 0$.
  - Depois aumentamos x_0 tal que a variável de folga mais negativa vire positiva.
  - x_0 é variável sainte x_k tal que $k = \operatorname{argmax}_{1\le i\le m} -b_i$.

Exemplo: Problema original

$$\begin{aligned} \textbf{maximiza} \quad & z = -2x_1 - x_2 \\ \textbf{sujeito a} \quad & -x_1 + x_2 \le -1 \\ & -x_1 - 2x_2 \le -2 \\ & x_2 \le 1 \\ & x_1, x_2 \ge 0 \end{aligned}$$

Exemplo: Problema auxiliar

$$\begin{aligned} \textbf{maximiza} \quad & z = -x_0 \\ \textbf{sujeito a} \quad & -x_1 + x_2 - x_0 \le -1 \\ & -x_1 - 2x_2 - x_0 \le -2 \\ & x_2 - x_0 \le 1 \\ & x_0, x_1, x_2 \ge 0 \end{aligned}$$

Exemplo: Dicionário inicial do problema auxiliar

$$
\begin{array}{rlllll}
 & & & & \downarrow \\
z & = & & & -x_0 \\
\hline
w_1 & = -1 & +x_1 & -x_2 & +x_0 \\
\leftarrow w_2 & = -2 & +x_1 & +2x_2 & +x_0 \\
w_3 & = 1 & & -x_2 & +x_0
\end{array}
$$

- Observe que a solução básica não é viável.
- Para achar uma solução básica viável: fazemos um primeiro pivot com variável entrante x_0 e variável sainte w_2.

Exemplo: Dicionário inicial viável do sistema auxiliar

$$
\begin{array}{rllll}
 & & & \downarrow & \\
z & = -2 & +x_1 & +2x_2 & -w_2 \\
\hline
\leftarrow w_1 & = 1 & & -3x_2 & +w_2 \\
x_0 & = 2 & -x_1 & -2x_2 & +w_2 \\
w_3 & = 3 & -x_1 & -3x_2 & +w_2
\end{array}
$$

Primeiro pivot

$$
\begin{array}{rllll}
 & & \downarrow & & \\
z & = -4/3 & +x_1 & -2/3w_1 & -1/3w_2 \\
\hline
x_2 & = 1/3 & & -1/3w_1 & +1/3w_2 \\
\leftarrow x_0 & = 4/3 & -x_1 & +2/3w_1 & +1/3w_2 \\
w_3 & = 2 & -x_1 & +w_1 &
\end{array}
$$

Segundo pivot

$$
\begin{array}{rllll}
z & = 0 & -x_0 & & \\
\hline
x_2 & = 1/3 & & -1/3w_1 & +1/3w_2 \\
x_1 & = 4/3 & -x_0 & +2/3w_1 & +1/3w_2 \\
w_3 & = 2/3 & +x_0 & +1/3w_1 & -1/3w_2
\end{array}
$$

Solução ótima!

Solução do sistema auxiliar

- O que vale a solução do sistema auxiliar?
- Obviamente, se o sistema original tem solução, o sistema auxiliar também tem uma solução com $x_0 = 0$.
- Logo, após aplicar o método Simplex ao sistema auxiliar, temos os casos
  - $x_0 > 0$: O sistema original não tem solução.
  - $x_0 = 0$: O sistema original tem solução. Podemos descartar x_0 e continuar resolvendo o sistema original com a solução básica viável obtida.
- A solução do sistema auxiliar se chama *fase I*, a solução do sistema original *fase II*.

Sistema original

Reescreve-se a função objetivo original substituindo as variáveis básicas do sistema original pelas equações correspondentes do sistema auxiliar, de forma que a função objetivo z não contenha variáveis básicas. No exemplo, a função objetivo é rescrita como:

$$z = -2x_1 - x_2 = -3 - w_1 - w_2.$$

$$
\begin{array}{rlll}
z & = -3 & -w_1 & -w_2 \\
\hline
x_2 & = 1/3 & -1/3w_1 & +1/3w_2 \\
x_1 & = 4/3 & +2/3w_1 & +1/3w_2 \\
w_3 & = 2/3 & +1/3w_1 & -1/3w_2
\end{array}
$$

Nesse exemplo, o dicionário original já é ótimo!

2.6 Soluções degeneradas

Solução degenerada

- Um dicionário é *degenerado* se existe um $\bar{b}_i = 0$.
- Qual o problema?
- Pode acontecer um pivot que não aumenta a variável entrante, e portanto não aumenta o valor da função objetivo.

Exemplo 1

- Nem sempre é um problema.

$$
\begin{array}{rlll}
 & & \downarrow & \\
z & = 5 & +x_3 & -x_1 \\
\hline
x_2 & = 5 & -2x_3 & -3x_1 \\
x_1 & = 7 & & -4x_1 \\
\leftarrow w_3 & = & & +x_1
\end{array}
$$

- x_2 é a variável sainte e o valor da função objetivo aumenta.

Exemplo 2

$$
\begin{array}{rllll}
 & & & \downarrow & \\
z & = 3 & -1/2x_1 & +2x_2 & -3/2w_1 \\
\hline
x_3 & = 1 & -1/2x_1 & & -1/2w_1 \\
w_2 & = & x_1 & -x_2 & +w_1
\end{array}
$$

- Se a variável sainte é determinada pela equação com $\bar{b}_i = 0$, temos um *pivot degenerado*.
- Nesse caso, a variável entrante não aumenta: temos a mesma solução depois do pivot.

Exemplo 2: Primeiro pivot

- Pivô: x_2–w_2

$$
\begin{array}{rllll}
 & & \downarrow & & \\
z & = 3 & +3/2x_1 & -2w_2 & +1/2w_1 \\
\hline
\leftarrow x_3 & = 1 & -1/2x_1 & & -1/2w_1 \\
x_2 & = & x_1 & -w_2 & +w_1
\end{array}
$$

- O valor da função objetivo não aumentou!

Exemplo 2: Segundo pivot

- Pivô: x_1–x_3

$$
\begin{array}{rllll}
z & = 6 & -3x_3 & -2w_2 & -w_1 \\
\hline
x_1 & = 2 & -2x_3 & -w_1 & \\
x_2 & = 2 & -2x_3 & -w_2 &
\end{array}
$$

- A segunda iteração aumentou o valor da função objetivo!

Ciclos

- O pior caso seria, se entramos em ciclos.
- É possível? Depende da regra de seleção de variáveis entrantes e saintes.
- Nossas regras
  - Escolha a variável entrante com o maior coeficiente.
  - Escolha a variável sainte que restringe mais.
  - Em caso de empate, escolha a variável com o menor índice.
- Ciclos são possíveis: O seguinte sistema possui um ciclo de 6 pivôs $x-1$–w_1, x_2–w_2, x_3–x_1, x_4–x_2, w_1–x_3, w_2–x_4.

$$
\begin{array}{rllllll}
z & = & 10x_1 & -57x_2 & -9x_3 & -24x_4 \\
\hline
w_1 & = & -1/2x_1 & +11/2x_2 & +5/2x_3 & -9x_4 \\
w_2 & = & -1/2x_1 & +3/2x_2 & +1/2x_3 & -x_4 \\
w_3 & = 1 & -x_1 & & &
\end{array}
$$

Soluções do problema

- Como resolver o problema?
- Duas propostas
  - Método lexicográfico.
  - Regra de Bland.

Método lexicográfico

- Idéia: O fato que existe um $\bar{b}_i = 0$ é por ocaso.
- Se introduzimos uma pequena perturbação $\epsilon \ll 1$
  - o problema desaparece
  - a solução será (praticamente) a mesma.

Método lexicográfico

- Ainda é possível que duas perturbações numéricas se cancelem.
- Para evitar isso: Trabalha-se simbolicamente.
- Introduzimos perturbações simbólicas

  $$0 < \epsilon_1 \ll \epsilon_2 \ll \cdots \ll \epsilon_m$$

  em cada equação.
- Característica: Todo ϵ_i é numa escala diferente dos outros tal que eles nunca se ancelem.

Exemplo

Sistema original degenerado e sistema perturbado

z	$=4$	$+2x_1$	$-x_2$
$1-4w_1$	$=1/2$		$-x_2$
w_2	$=$	$-2x_1$	$+4x_2$
w_3	$=$	x_1	$-3x_2$

z	$=4$				$+2x_1$	$-x_2$
w_1	$=1/2$	$+\epsilon_1$				$-x_2$
w_2	$=$		ϵ_2		$-2x_1$	$+4x_2$
w_3	$=$			ϵ_3	$+x_1$	$-3x_2$

Características

- Depois de chegar no valor ótimo, podemos retirar as perturbações ϵ_i.

  Teorema 2.1
  O método Simplex sempre termina escolhendo as variáveis saintes usando a regra lexicográfica.

Prova. É suficiente mostrar que o sistema nunca vai ser degenerado: assim o valor da função objetivo sempre cresce, e o método Simplex não entra em ciclo. A matriz de perturbações

$$\begin{pmatrix} \epsilon_1 & & & \\ & \epsilon_2 & & \\ & & \ldots & \\ & & & \epsilon_m \end{pmatrix}$$

inicialmente tem posto m (*rank* de uma matriz). As operações do método Simplex são operações lineares que não mudam o posto do matriz. Logo, em cada passo do método Simplex temos uma matriz de perturbações

$$\begin{pmatrix} e_{11}\epsilon_1 & e_{12} & \cdots & e_{1m}\epsilon_m \\ e_{21}\epsilon_1 & e_{22}\epsilon_2 & \cdots & e_{2m}\epsilon_m \\ \cdots & & \cdots & \\ e_{m1}\epsilon_1 & e_{m2}\epsilon_2 & \cdots & e_{mm}\epsilon_m \end{pmatrix}$$

que ainda tem posto m. Portanto, em cada linha i existe ao menos um $e_{ij} \neq 0$ e assim uma perturbação diferente de zero. Logo, o sistema não é degenerado. ■

Regra de Bland

- Outra solução do problema: A regra de Bland.
- Escolhe como variável entrante e sainte sempre a variável com o menor índice (caso tiver mais que um candidato).

  Teorema 2.2
  O método Simplex sempre termina se as variáveis entrantes e saintes são escolhidas através da regra de Bland.

Prova. Prova por contradição: Suponha que tem uma sequência de dicionários que entra num ciclo $D_0, D_1, \ldots, D_{k-1}$ usando a regra do Bland. Nesse ciclo algumas variáveis, chamadas *inconstantes*, entram e saem novamente da base, outras ficam sempre na base ou não-básicas. Seja x_t a variável inconstante, com o maior índice. Sem perda de generalidade, seja x_t a variável sainte do primeiro dicionário D_0. Seja x_s a variável entrante no D_0. Observe que x_s também é inconstante e portanto $s < t$. Seja D^* o dicionário em que x_t entra na base. Temos a seguinte situação

$$\begin{array}{ccccccc} x_s \text{ entra} & & & & x_t \text{ entra} & & \\ \downarrow & & & & \downarrow & & \\ D_0, & D_1, & D_2, & \cdots & D^*, & \cdots & D_{k-1} \\ \downarrow & & & & & & \\ x_t \text{ sai} & & & & & & \end{array}$$

com os sistemas correspondentes

D_0 :

$$z = z_0 + \sum_{j \in \mathcal{N}} c_j x_j$$

$$x_i = b_i - \sum_{j \in \mathcal{N}} a_{ij} x_j \quad i \in \mathcal{B}$$

D^* :

$$z = z^* + \sum_{j \in \mathcal{N}^*} c_j^* x_j$$

$$x_i = b_i^* - \sum_{j \in \mathcal{N}^*} a_{ij}^* x_j \quad i \in \mathcal{B}^*$$

Como temos um ciclo, todas variáveis inconstantes são nulas e o valor da função objetivo é constante. Logo $z_0 = z^*$. Se aumentamos o valor do x_s, qual seria o novo valor da função objetivo?

$$\begin{aligned} x_s &= y \\ x_j &= 0 \qquad j \in \mathcal{N} \setminus \{s\} \\ x_i &= b_i - a_{is} y \qquad i \in \mathcal{B} \end{aligned}$$

Logo temos o novo valor

$$z = z_0 + c_s y \tag{2.4}$$

(no sistema D_1) mas também

$$z = z^* + \sum_{j \in \mathcal{N}^*} c_j^* x_j = z_0 + \sum_{j \in \mathcal{N}^*} c_j^* x_j.$$

Agora, podemos substituir as variáveis com índices em $\mathcal{N}^* \cap \mathcal{B}$ nessa equação. Para facilitar a notação, vamos definir $c_j^* := 0$ para $j \notin \mathcal{N}^*$, que permite substituir todas variáveis $x_j, j \in \mathcal{B}$ e assim obtemos

$$z = z_0 + \sum_{j \in [1, n+m]} c_j^* x_j = z_0 + c_s^* y + \sum_{j \in \mathcal{B}} c_j^* (b_j - a_{js} y). \tag{2.5}$$

Equações 2.4 e 2.5 representam o mesmo valor, portanto

$$\left(c_s - c_s^* + \sum_{j \in \mathcal{B}} c_j^* a_{js} \right) = \sum_{j \in \mathcal{B}} c_j^* b_j$$

e como essa igualdade deve ser correta para qualquer aumento y

$$c_s - c_s^* + \sum_{j \in \mathcal{B}} c_j^* a_{js} = 0.$$

Como x_s entra em D_0 temos $c_s > 0$. Também, como $s < t$, mas x_t entra em D^*, $c_s^* \le 0$. Logo,

$$\sum_{j \in \mathcal{B}} c_j^* a_{js} < 0$$

e deve existir um $r \in \mathcal{B}$ tal que $c_r^* a_{rs} < 0$. Isso tem uma série de consequências:

1. $c_r^* \neq 0$.
2. $r \in \mathcal{N}^*$, porque somente as variáveis não-básicas tem c_j^* em D^*.
3. x_r é inconstante, porque ela é básica em D_0, mas não-básica em D^*.
4. $r \le t$, porque t foi a variável inconstante com o maior índice.
5. $r < t$, porque $c_t^* a_{ts} > 0$: x_t entra em D^*, logo $c_t^* > 0$, e x_t sai em D_0, logo $a_{ts} > 0$.
6. $c_r^* \le 0$, senão r e não t entraria em D^* seguindo a regra de Bland.
7. $a_{rs} > 0$.
8. $b_r = 0$, porque x_r é inconstante, mas todos variáveis inconstantes são nulas no ciclo, e x_r é básica em D_0.

Os últimos dois itens mostram que x_r foi candidato ao sair em D_0 com índice $r < t$, uma contradição à regra de Bland. ■

Teorema fundamental

Teorema 2.3 (Teorema fundamental da programação linear)
Para qualquer programa linear temos:

1. Se não existe solução ótima, o problema é inviável ou ilimitado.
2. Se existe uma solução viável, existe uma solução básica viável.
3. Se existe uma solução ótima, existe uma solução ótima básica.

2.7 Complexidade do método Simplex

Complexidade pessimista

- Com a regra de Bland o método Simplex sempre termina.

- Com $n+m$ variáveis (de decisão e de folga) existem

$$\binom{n+m}{n} = \binom{n+m}{m}$$

  soluções básicas possíveis.

- Logo: No pior caso o método Simplex termina depois essa número de pivots.

Complexidade pessimista

- Com as equações linearmente independentes, temos

$$\binom{n+m}{m} \le \binom{2n}{n}$$

- e os limites (exercício 5.11)

$$\frac{1}{2n}2^{2n} \le \binom{2n}{n} \le 2^{2n}.$$

- Logo, o número de passo no pior caso é *exponencial* no tamanho da entrada.

3 Dualidade

3.1 Introdução

Visão global

- *Dualidade*: Cada programa linear tem um programa linear associado, chamado de *dual*.
- Para enfatizar a diferença, o programa linear original chama-se de *primal*.
- Programas lineares duais tem várias aplicações como
  - Estimar a convergência.
  - Análise de sensibilidade e re-otimização de sistemas.
  - Solução mais simples ou eficiente com o Método Simplex dual.
- O programa dual também tem uma interpretação relevante.

Introdução

- Considere o programa linear

$$
\begin{aligned}
\textbf{maximiza} \quad & z = 4x_1 + x_2 + 3x_3 \\
\textbf{sujeito a} \quad & x_1 + 4x_2 \le 1 \\
& 3x_1 - x_2 + x_3 \le 3 \\
x_1, x_2, x_3 \ge 0 &
\end{aligned}
$$

- Cada solução viável fornece um limite inferior para o valor máximo.

$$
\begin{aligned}
x_1 = x_2 = x_3 = 0 &\Rightarrow z = 0 \\
x_1 = 3, x_2 = x_3 = 0 &\Rightarrow z = 4
\end{aligned}
$$

- Qual a qualidade da solução atual?
- Não sabemos, sem limite superior.

Limites superiores

- Como obter um limite superior? Observe: $z = 4x_1 + x_2 + 3x + 3 \le 10x_1 + x_2 + 3x_3 \le 10$
- Podemos construir uma combinação linear das desigualdades, tal que o coeficiente de cada x_j ultrapasse o coeficiente da função objetivo.
- Nosso exemplo:

$$
\begin{aligned}
& (x_1 + 4x_2) + 3(3x_1 - x_2 + x_3) \le 1 + 3 \cdot 3 = 10 \\
\iff & 10x_1 + x_2 + 3x_3 \le 10
\end{aligned}
$$

- Como obter um limite superior para a função objetivo?
- Qual seria o menor limite superior que esse método fornece?

O menor limite superior

- Sejam $y_1, \ldots, y_n$ os coeficientes de cada linha. Observação: Eles devem ser ≥ 0 para manter a direção das desigualdades.
- Então queremos

$$
\begin{aligned}
\textbf{minimiza} \quad & \sum_i b_i y_i \\
\textbf{sujeito a} \quad & \sum_i a_{ij} y_i \ge c_j \qquad 1 \le j \le n \\
& y_i \ge 0
\end{aligned}
$$

- Isto é o *problema dual*.

Dualidade: Características

- Em notação matricial

$$
\begin{aligned}
\textbf{maximiza} \quad & c^t x & \qquad \textbf{minimiza} \quad & b^t y \\
\textbf{sujeito a} \quad & Ax \le b & \qquad \textbf{sujeito a} \quad & y^t A \ge c^t \\
& x \ge 0 & & y \ge 0
\end{aligned}
$$

- O primeiro se chama *primal* e o segundo *dual*.
- Eles usam os mesmos parâmetros c_j, a_{ij}, b_i.

O dual do dual

- Observação: O dual do dual é o primal.
- Forma normal do dual:

$$\begin{array}{rl} \textbf{maximiza} & -b^t y \\ \textbf{sujeito a} & -y^t A \le -c^t \\ & y \ge 0 \end{array} \quad = \quad \begin{array}{rl} \textbf{maximiza} & -b^t y \\ \textbf{sujeito a} & (-A^t)y \le -c \\ & y \ge 0 \end{array}$$

- Dual do dual

$$\begin{array}{rl} \textbf{minimiza} & -c^t x \\ \textbf{sujeito a} & x^t(-A^t) \ge -b^t \\ & x \ge 0 \end{array} \quad = \quad \begin{array}{rl} \textbf{maximiza} & c^t x \\ \textbf{sujeito a} & Ax \le b \\ & x \ge 0 \end{array}$$

3.2 Interpretação do dual

Exemplo: Dieta dual

- Problema da dieta: Minimiza custos de uma dieta x que alcance dados VDR mínimos.

$$\begin{array}{rl} \textbf{minimiza} & c^t x \\ \textbf{sujeito a} & Ax \ge r \\ & x \ge 0 \end{array}$$

- Unidades das variáveis e parâmetros
  - x: Quantidade do alimento $[g]$
  - c: R$/alimento $[R\$/g]$
  - a_{ij}: Nutriente/Alimento $[g/g]$
  - r: Quantidade de nutriente $[g]$.

Exemplo: Dieta dual

- O problema dual é

$$\begin{array}{rl} \textbf{maximiza} & y^t r \\ \textbf{sujeito a} & y^t A \le c^t \\ & y \ge 0 \end{array}$$

- Qual a unidade de y? Preço por nutriente $[R\$/g]$.
- Imagine uma empresa, que produz cápsulas que substituem os nutrientes.
- Para vender no mercado, a empresa tem que garantir que uma dieta baseado em cápsulas custa menos que os alimentos correspondentes:

$$\sum_i y_i a_{ij} \le c_j$$

- Além disso, ela define preços por nutriente que maximizam o custo de uma dieta adequada, para maximizar o próprio lucro.

$$\textbf{maximiza} \quad y^t r$$

Interpretação do dual

- Outra interpretação: o valor de uma variável dual y_j é o *lucro marginal* de adicionar mais uma unidade b_j.

  Teorema 3.1
  Se um sistema tem ao menos uma solução básica viável não-degenerada, existe um ϵ tal que, se $|t_j| \le \epsilon$ para $1 \le j \le m$,

$$\begin{array}{rl} \textbf{maximiza} & c^t x \\ \textbf{sujeito a} & Ax \le b + t \\ & x \ge 0 \end{array}$$

  tem uma solução ótima com valor

$$z = z^* + {y^*}^t t$$

  (com z^* o valor ótimo do primal, é y^* a solução ótima do dual).

3.3 Características

Teorema da dualidade fraca

Teorema 3.2 (Dualidade fraca)
Se $x_1, \ldots, x_n$ é uma solução viável do sistema primal, e $y_1, \ldots, y_m$ uma solução viável do sistema dual, então

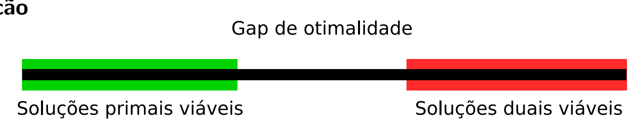
$$\sum_{1 \le i \le n} c_i x_i \le \sum_{1 \le j \le m} b_j y_j.$$

Prova.

$$
\begin{aligned}
& c^t x && && (3.1)\\
\le & (y^t A)x = y^t(Ax) && \text{pela restrição dual} && (3.2)\\
\le & y^t b && \text{pela restrição primal} && (3.3)
\end{aligned}
$$

■

Situação



- Em aberto: Qual o tamanho desse intervalo em geral?

Teorema da da dualidade forte

Teorema 3.3
Se $x_1^*, \dots, x_n^*$ é uma solução ótima do sistema primal, existe uma solução ótima $y_1^*, \dots, y_m^*$ do sistema dual, e

$$\sum_{1\le i\le n} c_i x_i^* = \sum_{1\le j\le m} b_j y_j^*.$$

Prova. Seja x^* uma solução ótima do sistema primal, que obtemos pelo método Simplex. No início introduzimos variáveis de folga

$$x_{n+j} = b_j - \sum_{1\le i\le n} a_{ji} x_i \qquad 1 \le j \le m$$

e a função objetivo final é

$$z = z^* + \sum_{1\le i\le n+m} \bar{c}_i x_i$$

(supondo que $\bar{c}_i = 0$ para variáveis básicas). Temos que construir uma solução ótima dual y^*. Pela optimalidade, na função objetivo acima, todos $\overline{c_i}$ devem ser não-positivos. Afirmamos que $y_j^* = -\overline{c_{n+j}} \ge 0$ para $j \in [1, m]$ é uma solução dual ótima. Como z^* o valor ótimo do problema inicial, temos $z^* = \sum_{1\le i\le n} c_i x_i^*$.

Reescrevendo a função objetivo temos

$$
\begin{aligned}
z & \\
&= \sum_{1\le i\le n} c_i x_i && \text{sistema inicial}\\
&= z^* + \sum_{1\le i\le n+m} \bar{c}_i x_i && \text{sistema final}\\
&= z^* + \sum_{1\le i\le n} \bar{c}_i x_i + \sum_{1\le j\le m} \overline{c_{n+j}} x_{n+j} && \text{separando índices}\\
&= z^* + \sum_{1\le i\le n} \bar{c}_i x_i - \sum_{1\le j\le m} y_j^* \left(b_j - \sum_{1\le i\le n} a_{ji} x_i \right) && \text{subst. solução e var. folga}\\
&= \left(z^* - \sum_{1\le j\le m} y_j^* b_j \right) + \sum_{1\le i\le n} \left(\bar{c}_i + \sum_{1\le j\le m} y_j^* a_{ji} \right) x_i && \text{agrupando}
\end{aligned}
$$

Essa derivação está válida para x_i qualquer, porque são duas expressões para a mesma função objetivo, portanto

$$z^* = \sum_{1\le j\le m} y_j^* b_j \quad \text{e} \quad c_i = \bar{c}_i + \sum_{1\le j\le m} y_j^* a_{ji} \qquad 1 \le i \le n.$$

Com isso sabemos que o primal e dual possuem o mesmo valor

$$\sum_{1\le j\le m} y_j^* b_j = z^* = \sum_{1\le i\le n} c_i x_i^*$$

e como $\bar{c}_i \le 0$ sabemos que a solução y^* satisfaz a restrições duais

$$
\begin{aligned}
c_i \le \sum_{1\le j\le m} y_j^* a_{ji} \qquad & 1 \le i \le n\\
y_i^* \ge 0 \qquad & 1 \le i \le m
\end{aligned}
$$

■

Consequências: Soluções primais e duais

- Com o teorema da dualidade forte, temos quatro possibilidades

Sistema primal	Sistema dual	Intervalo
Ótima	Ótima	Sem
Ilimitado	Inviável	Sem
Inviável	Ilimitado	Sem
Inviável	Inviável	Infinito

Exemplo 3.1
Pelo teorema da dualidade forte, não podemos concluir, que existe um caso que tanto o sistema primal qunato o sistema dual são inviáveis. O seguinte exemplo mostra que isso pode realmente acontecer. O sistema primal

$$\begin{array}{rl} \textbf{maximiza} & x_1 \\ \textbf{sujeito a} & +x_1 - x_2 \leq 0 \\ & -x_1 + x_2 \leq -1 \end{array}$$

possui sistema dual correspondete

$$\begin{array}{rl} \textbf{minimiza} & -y_2 \\ \textbf{sujeito a} & +y_1 - y_2 \geq 1 \\ & -y_1 + y_2 \geq 0 \end{array}$$

Os dois sistemas são inviáveis. ◊

Consequências

- Dado soluções primais e duais x^*, y^* tal que $c^t x^* = b^t y^*$ podemos concluir que ambas soluções são ótimas (x^*, y^* é um *certificado* da optimalidade)[1].
- A prova mostra: com o valor ótimo do sistema primal, sabemos também o valor ótima do sistema dual.
- Além disso: Podemos trocar livremente entre o sistema primal e dual. $\Rightarrow$ Método Simplex dual.

Outra consequência do Teorema da dualidade forte é o

Teorema 3.4 (Teorema das folgas complementares)
Se x^*, y^* são soluções ótimas do sistema primal e dual, respectivamente, temos

$$y^{*t}(b - Ax) = 0 \tag{3.4}$$
$$(y^{*t}A - c^t)x^* = 0 \tag{3.5}$$

[1] Uma consequencia é que o problema de decisão correspondente, determinar se existe uma solução maior que um dado valor, possui um certificado que pode ser verificado em tempo polinomial tanto para uma resposta positiva quanto uma resposta negativa. Portanto, já antes da descoberta de um algoritmo polinomial para esse problema, foi claro que ele pertence a NP ∩ co-NP.

Prova. Pelo Teorema da dualidade forte as duas desigualdades 3.2 e 3.3 da prova do Teorema da dualidade fraca se tornam igualdades para soluções ótimas:

$$c^t x^* = y^{*t} A x^* = y^{*t} b$$

Reagrupando termos, o teorema segue. ■
As igualdades 3.4 e 3.5 são ainda válidas em cada componente, porque tanto as soluções ótimas x^*, y^* quanto as folgas primas e duais $b - Ax$ e $y^{*t}A - c^t$ sempre são positivos.

$$x_i > 0 \Rightarrow \sum_{1 \leq j \leq m} y_j a_{ji} = c_i \tag{3.6}$$
$$\sum_{1 \leq j \leq m} y_j a_{ji} > c_i \Rightarrow x_i = 0 \tag{3.7}$$
$$y_j > 0 \Rightarrow b_j = \sum_{1 \leq i \leq n} a_{ji} x_i \tag{3.8}$$
$$b_j > \sum_{1 \leq i \leq n} a_{ji} x_i \Rightarrow y_j = 0 \tag{3.9}$$

Como consequência, podemos ver que, por exemplo, caso uma igualdade primal não possui folga, a variável dual correspondente é positiva, e, contrariamente, caso uma igualdade primal possui folga, a variável dual correspondente é zero. As mesmas relações se aplicam para as desigualdades no sistema dual. Após a introdução da forma matricial no seção 3.6 vamos analisar a interpretação das variáveis duais com mais detalha no seção 3.7.

3.4 Método Simplex dual

Método Simplex dual

- Considere

$$\begin{array}{rl} \textbf{maximiza} & -x_1 - x_2 \\ \textbf{sujeito a} & -2x_1 - x_2 \leq 4 \\ & -2x_1 + 4x_2 \leq -8 \\ & -x_1 + 3x_2 \leq -7 \\ & x_1, x_2 \geq 0 \end{array}$$

- Qual o dual?

$$\begin{array}{rl} \textbf{minimiza} & 4y_1 - 8y_2 - 7y_3 \\ \textbf{sujeito a} & -2y_1 - 2y_2 - y_3 \geq -1 \\ & -y_1 + 4y_2 + 3y_2 \geq -1 \end{array}$$

Com dicionários

$$\begin{array}{llll}
z & = & -x_1 & -x_2 \\ \hline
w_1 & = 4 & +2x_1 & +x_2 \\
w_2 & = -8 & +2x_1 & -4x_2 \\
w_3 & = -7 & +x_1 & -3x_2
\end{array}$$

$$\begin{array}{lllll}
-w & = & -4y_1 & +8y_2 & +7y_3 \\ \hline
z_1 & = 1 & -2y_1 & -2\mathbf{y_2} & -y_3 \\
z_2 & = 1 & -y_1 & +4y_2 & +3y_3
\end{array}$$

- Observação: O primal não é viável, mas o dual é!
- Correspondência das variáveis:

	Variáveis	
	principais	de folga
Primal	$x_1, \ldots, x_n$	$w_1, \ldots, w_m$
Dual	$z_1, \ldots, z_n,$ de folga	$y_1, \ldots, y_m$ principais

- Primeiro pivot: y_2 entra, z_1 sai. No primal: w_2 sai, x_1 entra.

Primeiro pivot

$$\begin{array}{llll}
z & = -4 & -0.5w_2 & -3x_2 \\ \hline
w_1 & = 12 & +w_2 & +5x_2 \\
x_1 & = 4 & +0.5w_2 & +2x_2 \\
w_3 & = -3 & +0.5w_2 & -x_2
\end{array}$$

$$\begin{array}{lllll}
-w & = 4 & -12y_1 & -4z_1 & +3y_3 \\ \hline
y_2 & = 0.5 & -y_1 & -0.5z_1 & -0.5\mathbf{y_3} \\
z_2 & = 3 & -5y_1 & -2z_1 & +y_3
\end{array}$$

- Segundo pivot: y_3 entra, y_2 sai. No primal: w_3 sai, w_2 entra.

Segundo pivot

$$\begin{array}{llll}
z & = -7 & -w_3 & -4x_2 \\ \hline
w_1 & = 18 & +2w_3 & +7x_2 \\
x_1 & = 7 & +w_3 & +3x_2 \\
w_2 & = 6 & +2w_3 & +2x_2
\end{array}$$

$$\begin{array}{lllll}
-w & = 7 & -18y_1 & -7z_1 & -6y_2 \\ \hline
y_3 & = 1 & -2y_1 & -z_1 & -2y_2 \\
z_2 & = 4 & -7y_1 & -3z_1 & -2y_2
\end{array}$$

- Sistema dual é ótimo, e portanto o sistema primal também.

Método Simplex dual

- Observação: Não é necessário escrever o sistema dual. Ele é sempre o negativo transposto do sistema primal.

$$z = \bar{z} + \sum_{j \in \mathcal{N}} \bar{c}_j x_j$$
$$x_i = \bar{b}_i - \sum_{j \in \mathcal{N}} \bar{a}_{ij} x_j \qquad i \in \mathcal{B}$$

- Mas é necessário modificar as regras para resolver o sistema dual.

Método Simplex dual: Viabilidade e otimalidade

- Pré-condição: O dicionário é *dualmente viável*, i.e. os coeficientes das variáveis não-básicas na função objetivo tem quer ser não-positivos.

$$\bar{c}_j \leq 0 \quad \text{para} \quad j \in \mathcal{N}.$$

- Otimalidade: Todos variáveis básicas primais positivas

$$\forall i \in \mathcal{B} : \bar{b}_i \geq 0$$

Método Simplex dual: Pivô

- Caso existe uma variável primal negativa: A solução dual não é ótima.
- Regra do maior coeficiente: A variável básica primal com menor valor (que é negativo) sai da base primal.

$$i = \operatorname*{argmin}_{i \in \mathcal{B}} \bar{b}_i$$

- A variável primal nula com fração $\bar{a}_{ij}/\bar{c}_j$ maior entra.

$$j = \operatorname*{argmin}_{\substack{j \in \mathcal{N} \\ \bar{a}_{ij} < 0}} \frac{\bar{c}_j}{\bar{a}_{ij}} = \operatorname*{argmax}_{\substack{j \in \mathcal{N} \\ \bar{a}_{ij} < 0}} \frac{\bar{a}_{ij}}{\bar{c}_j} = \operatorname*{argmax}_{j \in \mathcal{N}} \frac{\bar{a}_{ij}}{\bar{c}_j}$$

Método Simplex dual

Resumo:

- Dualmente viável: $\bar{c}_j \leq 0$ para $j \in \mathcal{N}$.
- Otimalidade: $\forall i \in \mathcal{B} : \bar{b}_i \geq 0$.
- Variável sainte: $i = \operatorname{argmin}_{i \in \mathcal{B}} \bar{b}_i$
- Variável entrante: $j = \operatorname{argmax}_{j \in \mathcal{N}} \frac{\bar{a}_{ij}}{\bar{c}_j}$.

Exemplo

$$\begin{array}{rl} \textbf{maximiza} & z = -2x_1 - x_2 \\ \textbf{sujeito a} & -x_1 + x_2 \leq -1 \\ & -x_1 - 2x_2 \leq -2 \\ & x_2 \leq 1 \\ & x_1, x_2 \geq 0 \end{array}$$

Exemplo: Dicionário inicial

$$\begin{array}{lrrr} z & = & -2x_1 & -x_2 \\ w_1 & = -1 & +x_1 & -x_2 \\ w_2 & = -2 & +x_1 & +2\mathbf{x_2} \\ w_3 & = 1 & & -x_2 \end{array}$$

- O dicionário primal não é viável, mais o dual é.

Exemplo: Primeiro pivot

$$\begin{array}{lrrr} z & = -1 & -3/2x_1 & -1/2w_2 \\ w_1 & = -2 & +3/2\mathbf{x_1} & -1/2w_2 \\ x_2 & = 1 & -1/2x_1 & +1/2w_2 \\ w_3 & = & +1/2x_1 & -1/2w_2 \end{array}$$

Exemplo: Terceiro privot

$$\begin{array}{lrrr} z & = -3 & -w_1 & -w_2 \\ x_1 & = 4/3 & +2/3w_1 & +1/3w_2 \\ x_2 & = 1/3 & -1/3w_1 & +1/3w_2 \\ w_3 & = 2/3 & +1/3w_1 & -1/3w_2 \end{array}$$

3.5 Dualidade em forma não-padrão

Dualidade em forma padrão

$$\begin{array}{rl} \textbf{maximiza} & c^t x \\ \textbf{sujeito a} & Ax \leq b \\ & x \geq 0 \end{array} \qquad \begin{array}{rl} \textbf{minimiza} & b^t y \\ \textbf{sujeito a} & y^t A \geq c^t \\ & y \geq 0 \end{array}$$

- O que acontece se o sistema não é em forma padrão?

Igualdades

- Caso de igualdades: Substituindo desigualdades..

$$\begin{array}{rl} \textbf{maximiza} & c^t x \\ \textbf{sujeito a} & Ax = b \\ & x \geq 0 \end{array} \qquad \begin{array}{rl} \textbf{maximiza} & c^t x \\ \textbf{sujeito a} & Ax \leq b \\ & Ax \geq b \\ & x \geq 0 \end{array}$$

- ... padronizar novamente, e formar o dual:

$$\begin{array}{rl} \textbf{maximiza} & c^t x \\ \textbf{sujeito a} & Ax \leq b \\ & -Ax \leq -b \\ & x \geq 0 \end{array} \qquad \begin{array}{rl} \textbf{minimiza} & b^t y^+ - b^t y^- \\ \textbf{sujeito a} & {y^+}^t A - {y^-}^t A \geq c \\ & y^+ \geq 0, y^- \geq 0 \\ & y^+ = (y_1^+, \ldots, y_m^+)^t \\ & y^- = (y_1^-, \ldots, y_m^-)^t \end{array}$$

Igualdades

- Equivalente, usando variáveis não-restritas $y = y^+ - y^-$

$$\begin{array}{rl} \textbf{minimiza} & b^t y \\ \textbf{sujeito a} & y^t A \geq c \\ & y^t \lesseqgtr 0 \end{array}$$

- Resumo

Primal	Dual
Igualdade	Variável dual livre
Desigualdade ($\leq$)	Variável dual não-negativa
Variável primal livre	Igualdade
Variável primal não-negativa	Desigualdade ($\geq$)

3.6 Os métodos em forma matricial

A forma matricial permite uma descrição mais compacto do método Simplex. A seguir vamos resumir os métodos Simplex primal e dual na forma matricial. Mais importante, nesse forma é possível expressar o dicionário correspondente com qualquer base em termos das dados inicias (A, c, b). Na próxima seção vamos usar essa forma para analisar a sensibilidade de uma solução ao pequenas perturbações dos dados (i.e. os coeficientes A,b, e c).

Sistema padrão

- O sistema padrão é

$$\begin{aligned} \textbf{maximiza} \quad & c^t x \\ \textbf{sujeito a} \quad & Ax \leq b \\ & x \geq 0 \end{aligned}$$

- Com variáveis de folga $x_{n+1}, \ldots, x_{n+m}$ e A,c,x *novo* (definição segue abaixo)

$$\begin{aligned} \textbf{maximiza} \quad & c^t x \\ \textbf{sujeito a} \quad & Ax = b \\ & x \geq 0 \end{aligned}$$

Matrizes

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} & 1 & & & \\ a_{21} & a_{22} & \cdots & a_{2n} & & 1 & & \\ \vdots & \vdots & & \vdots & & & \ddots & \\ a_{m1} & a_{m2} & \ldots & a_{mn} & & & & 1 \end{pmatrix};$$

$$b = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}; c = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \\ 0 \\ \vdots \\ 0 \end{pmatrix}; x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \\ x_{n+1} \\ \vdots \\ x_{n+m} \end{pmatrix}$$

Separação das variáveis

- Em cada iteração as variáveis estão separados em básicas e não-básicas.
- Conjuntos de índices correspondentes: $\mathcal{B} \dot\cup \mathcal{N} = [1, n+m]$.
- A componente i de Ax pode ser separado como

$$\sum_{1 \leq j \leq n+m} a_{ij} x_j = \sum_{j \in \mathcal{B}} a_{ij} x_j + \sum_{j \in \mathcal{N}} a_{ij} x_j$$

Separação das variáveis

- Para obter a mesma separação na forma matricial: Reordenamos as colunas e separamos as matrizes e vetores:

$$A = (B\ N)\,; x = \begin{pmatrix} x_B \\ x_N \end{pmatrix}; c = \begin{pmatrix} c_B \\ c_N \end{pmatrix}$$

- com $B \in \mathbb{R}^{m \times m}$, $N \in \mathbb{R}^{m \times n}$, $c \in \mathbb{R}^{n+m}$.

Forma matricial das equações

- Agora, $Ax = b$ é equivalente com

$$(B\ N) \begin{pmatrix} x_B \\ x_N \end{pmatrix} = Bx_B + Nx_N = b$$

- Numa solução básica, a matriz B tem posto m tal que as colunas de B formam uma *base* do $\mathbb{R}^m$. Logo B tem inversa e

$$x_B = B^{-1}(b - Nx_N) = B^{-1}b - B^{-1}Nx_N$$

Forma matricial da função objetivo

- A função objetivo é

$$z = c^t x = (c_B\ c_N)\begin{pmatrix} x_B \\ x_N \end{pmatrix} = c_B^t x_B + c_N^t x_N$$

- e usando $x_B = B^{-1}b - B^{-1}Nx_N$ obtemos

$$\begin{aligned} z &= c_B^t(B^{-1}b - B^{-1}Nx_N) + c_N^t x_N \\ &= c_B^t B^{-1}b - (c_B^t B^{-1}N - c_N^t)x_N \\ &= c_B^t B^{-1}b - ((B^{-1}N)^t c_B - c_N)^t x_N \end{aligned}$$

Dicionário em forma matricial

- Logo, o dicionário em forma matricial é

$$\begin{aligned} z &= c_B^t B^{-1}b - ((B^{-1}N)^t c_B - c_N)^t x_N \\ x_B &= B^{-1}b - B^{-1}Nx_N \end{aligned}$$

- Compare com a forma em componentes:

$$\begin{aligned} z &= \bar{z} + \sum_{j\in\mathcal{N}} \bar{c}_j x_j \\ x_i &= \bar{b}_i - \sum_{j\in\mathcal{N}} \bar{a}_{ij} x_j \qquad i \in \mathcal{B} \end{aligned}$$

Dicionário em forma matricial

- Portanto, vamos identificar

$$\begin{aligned} \bar{z} &= c_B^t B^{-1}b; & \bar{c} &= -((B^{-1}N)^t c_B - c_N)^t \\ \bar{b} &= B^{-1}b; & \bar{A} &= (\bar{a}_{ij}) = B^{-1}N \end{aligned}$$

- para obter o dicionário

$$\begin{aligned} z &= \bar{z} + \bar{c}^t x_N \\ x_B &= \bar{b} - \bar{A}x_N \end{aligned}$$

Sistema dual

- As variáveis primais são

$$x = (\underbrace{x_1 \ldots x_n}_{\text{original}}\ \underbrace{x_{n+1} \ldots x_{n+m}}_{\text{folga}})^t$$

- Para manter índices correspondentes, escolhemos variáveis duais da forma

$$y = (\underbrace{y_1 \ldots y_n}_{\text{folga}}\ \underbrace{y_{n+1} \ldots y_{n+m}}_{\text{dual}})^t$$

- O dicionário do dual correspondente então é

Primal	Dual
$z = \bar{z} + \bar{c}^t x_N$	$-w = -\bar{z} - \bar{b}^t y_B$
$x_B = \bar{b} - \bar{A}x_N$	$y_N = -\bar{c} + \bar{A}^t y_B$

Primal e dual

- A solução básica do sistema primal é

$$x_N^* = 0; \quad x_B^* = \bar{b} = B^{-1}b$$

- A solução dual correspondente é

$$y_B^* = 0; \quad y_N^* = -\bar{c} = ((B^{-1}N)^t c_B - c_N)^t$$

- Com isso temos os dicionários

$$\begin{aligned} z &= \bar{z} - (y_N^*)^t x_N & -w &= -\bar{z} - (x_B^*)^t y_B \\ x_B &= x_B^* - (B^{-1}N)x_N & y_N &= y_N^* + (B^{-1}N)^t y_B \end{aligned}$$

Método Simplex em forma matricial

- Começamos com uma partição $\mathcal{B} \dot\cup \mathcal{N} = [1, n+m]$.
- Em cada iteração selecionamos uma variável sainte $i \in \mathcal{B}$ e entrante $j \in \mathcal{N}$.
- Fazemos o pivot x_i com x_j.
- Depois a nova base é $\mathcal{B} \setminus \{i\} \cup \{j\}$.

Método Simplex em forma matricial

S1: Verifique solução ótima Se $y_N^* \geq 0$ a solução atual é ótima. Pare.

S2: Escolhe variável entrante Escolhe $j \in \mathcal{N}$ com $y_j^* < 0$. x_j é a variável entrante.

S3: Determine passo básico Aumentando x_j uma unidade temos novas variáveis não-básicas $x_N = x_N^* + \Delta x_N$ com $\Delta x_N = (0 \cdots 010 \cdots 0)^t = e_j$ e e_j o vetor nulo com somente 1 na posição correspondente com índice j. Como

$$x_B = x_B^* - B^{-1} N x_N$$

a diminuição correspondente das variáveis básicas é $\Delta x_B = B^{-1} N e_j$.

Método Simplex em forma matricial

S4: Determine aumento máximo O aumento máximo de x_j é limitado por $x_B \geq 0$, i.e.

$$x_B = x_B^* - t\Delta x_B \geq 0 \iff x_B^* \geq t\Delta x_B.$$

Com $t, x_B^* \geq 0$ temos

$$t \leq t^* = \min_{\substack{i \in \mathcal{B} \\ \Delta x_i > 0}} \frac{x_i^*}{\Delta x_i}$$

S5: Escolhe variável sainte Escolhe um $i \in \mathcal{B}$ com $x_i^* = t^* \Delta x_i$.

Método Simplex em forma matricial

S5: Determine passo dual A variável entrante dual é y_i. Aumentando uma unidade, as variáveis y_N diminuem $\Delta y_N = -(B^{-1}N)^t e_i$.

S6: Determina aumento máximo Com variável sainte y_j, sabemos que y_i pode aumentar ao máximo

$$s = \frac{y_j^*}{\Delta y_j}.$$

S7: Atualiza solução

$$\begin{aligned} x_j^* &:= t & \qquad y_i^* &:= s \\ x_B^* &:= x_B^* - t\Delta x_B & \qquad y_N^* &:= y_N^* - s\Delta y_N \\ \mathcal{B} &:= \mathcal{B} \setminus \{i\} \cup \{j\} \end{aligned}$$

3.7 Análise de sensibilidade

Motivação

- Tratamos os parâmetros como ser fixados.
- Qual o efeito de uma perturbação

$$c := c + \Delta c; \quad b := b + \Delta b; \quad A := A + \Delta A?$$

(Imagina erros de medida, pequenas flutuações, etc.)

Análise de sensibilidade

- Após a solução de um sistema linear, temos o dicionário ótimo

$$\begin{aligned} z &= z^* - (y_N^*)^t x_N \\ x_B &= x_B^* - B^{-1} N x_N \end{aligned}$$

- com

$$\begin{aligned} x_B^* &= B^{-1} b \\ y_N^* &= ((B^{-1}N)^t c_B - c_N) \\ z^* &= c_B^t B^{-1} b \end{aligned}$$

Modificar c

- Mudarmos c para $\hat{c}$, mantendo a base $\mathcal{B}$.
- x_B^* não muda, mas temos que reavaliar y_N^* e z^*.
- Depois, x_B^* ainda é uma solução básica viável do sistema primal.
- Logo, podemos continuar aplicando o método Simplex primal.

Modificar b

- Da mesma forma, modificamos b para $\hat{b}$ (mantendo a base).
- y_N^* não muda, mas temos que reavaliar x_B^* e z^*.
- Depois, y_N^* ainda é uma solução básica viável do sistema dual.
- Logo, podemos continuar aplicando o método Simplex dual.

Vantagem dessa abordagem

- Nos dois casos, esperamos que a solução inicial já é perto da solução ótima.
- Experiência prática confirma isso.
- O que acontece se queremos modificar tanto b quanto c ou ainda A?
- A solução atual não necessariamente é viável no sistema primal ou dual.
- Mas: Mesmo assim, a convergência na prática é mais rápido.

Estimar intervalos

- Pergunta estendida: Qual o intervalo de $t \in R$ tal que o sistema com $\hat{c} = c + t\Delta c$ permanece ótimo?
- Para $t = 1$: $y_N^* = ((B^{-1}N)^t c_B - c_N)$ aumenta $\Delta y_N := (B^{-1}N)^t \Delta c_B - \Delta c_N$.
- Em geral: Aumento $t\Delta y_N$.
- Condição para manter a viabilidade dual:

$$y_N^* + t\Delta y_N \geq 0$$

- Para $t > 0$ temos

$$t \leq \min_{\substack{j \in \mathcal{N} \\ \Delta y_j < 0}} -\frac{y_j^*}{\Delta y_j}$$

- Para $t < 0$ temos

$$\max_{\substack{j \in \mathcal{N} \\ \Delta y_j > 0}} -\frac{y_j^*}{\Delta y_j} \leq t$$

Estimar intervalos

- Agora seja $\hat{b} = b + t\Delta b$.
- Para $t = 1$: $x_B^* = B^{-1}b$ aumenta $\Delta x_B := B^{-1}\Delta b$.
- Em geral: Aumento $t\Delta b$.

- Condição para manter a viabilidade primal:

$$x_B^* + t\Delta x_B \geq 0$$

- Para $t > 0$ temos

$$t \leq \min_{\substack{i \in \mathcal{B} \\ \Delta x_i < 0}} -\frac{x_i^*}{\Delta x_i}$$

- Para $t < 0$ temos

$$\max_{\substack{i \in \mathcal{B} \\ \Delta x_i > 0}} -\frac{x_i^*}{\Delta x_i} \leq t$$

Exemplo 3.2
Considere o problema da empresa de aço (vista na aula prática, veja também execício 5.6).

$$\begin{aligned} \textbf{maximiza} \quad & 25p + 30c \\ \textbf{sujeito a} \quad & 7p + 10c \leq 56000 \\ & p \leq 6000 \\ & c \leq 4000 \end{aligned}$$

Qual o intervalo em que o valor do lucro das placas de $25R\$$ pode variar sem alterar a solução ótima?

Exemplo: Empresa de aço

- Sistema ótimo

$$\begin{array}{rlll} z & = 192000 & -3w_1 & -4w_2 \\ \hline p & = 6000 & & -w_2 \\ w_3 & = 2600 & +1/10w_1 & -7/10w_2 \\ c & = 1400 & -1/10w_1 & +7/10w_2 \end{array} \qquad \begin{array}{rll} z & = \bar{z} & +\bar{c}x_N \\ \hline x_B & = B^{-1}b & -B^{-1}Nx_N \end{array}$$

- Base $\mathcal{B} = \{p, w_3, c\}$, variáveis não-básicas $\mathcal{N} = \{w_1, w_2\}$. (Observe: Usamos conjuntos de variáveis, ao invés de conjuntos de índices).

Exemplo: Variáveis

- Vetores c e Δc. Observe que reordenamos dos dados do sistema inicial de forma correspondente com a ordem das variáveis do sistema final.

$$c = \begin{pmatrix} 25 \\ 0 \\ 30 \\ 0 \\ 0 \end{pmatrix}; c_B = \begin{pmatrix} 25 \\ 0 \\ 30 \end{pmatrix}; c_N = \begin{pmatrix} 0 \\ 0 \end{pmatrix};$$

$$\Delta c = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}; \Delta c_B = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}; \Delta c_N = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

Exemplo: Aumentos

- Aumento das variáveis duais

$$\Delta y_N = (B^{-1}N)^t \Delta c_B - \Delta c_N = (B^{-1}N)^t \Delta c_B$$

- com

$$B^{-1}N = \begin{pmatrix} 0 & 1 \\ -1/10 & 7/10 \\ 1/10 & -7/10 \end{pmatrix}$$

- temos

$$\Delta y_N = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

Exemplo: Limites

- Limites em geral

$$\max_{\substack{j \in \mathcal{N} \\ \Delta y_j > 0}} -\frac{y_j^*}{\Delta y_j} \leq t \leq \min_{\substack{j \in \mathcal{N} \\ \Delta y_j < 0}} -\frac{y_j^*}{\Delta y_j}$$

- Logo

$$-4 \leq t \leq \infty.$$

- Uma variação do preço entre $25 + [-4, \infty] = [21, \infty]$ preserve a otimalidade da solução atual.

- O valor da função objetivo pode variar, mas os valores das variáveis p e c permanecem os mesmos.

$\Diamond$

Exemplo 3.3
Qual o intervalo em que o lucro das placas (R$ 25) e dos canos (R$ 30) podem variar sem que a solução ótima seja alterada?

Exemplo: Variação do lucro dos placas e canos

- Neste caso, os vetores c, c_B, c_N e Δc_N permanecem os mesmos do exemplo anterior. Enquanto que:

$$\Delta c = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}; \Delta c_B = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix};$$

- Neste caso, o valor de Δy_N é

$$\Delta y_N = (B^{-1}N)^t \Delta c_B = \begin{pmatrix} 0 & -1/10 & 1/10 \\ 1 & 7/10 & -7/10 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1/10 \\ 3/10 \end{pmatrix};$$

- Logo $-40/3 \leq t \leq \infty$
- Ou seja, uma variação do lucro das placas entre R$ 11.67 e ∞, e do lucro dos canos entre R$ 16.67 e ∞, não altera a solução ótima do sistema.

$\Diamond$

Exemplo: Modificação

- Qual o intervalo em que o lucro dos canos (R$ 30) podem variar sem que a solução ótima seja alterada?
- Neste caso, os vetores c, c_B, c_N e Δc_N permanecem os mesmos do exemplo anterior. Enquanto que:

$$\Delta c = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}; \Delta c_B = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix};$$

- Neste caso, o valor de Δy_N é:

$$\Delta c_B = \begin{pmatrix} 1/10 \\ -7/10 \end{pmatrix};$$

- Logo $-30 \le t \le 40/7$
- Ou seja, uma variação do lucro dos canos entre R\$ 0 e R\$ 35.71, não altera a solução ótima do sistema.

Exemplo 3.4
O que acontece se mudarmos o lucro das placas para R\$ 20?

Exemplo: Placas com lucro R\$ 20

- Novos vetores

$$\hat{c} = \begin{pmatrix} 20 \\ 0 \\ 30 \\ 0 \\ 0 \end{pmatrix}; \hat{c}_B = \begin{pmatrix} 20 \\ 0 \\ 30 \end{pmatrix}; \hat{c}_N = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

- Aumento

$$\begin{aligned} \hat{y}_N^* &= (B^{-1}N)^t \hat{c}_B - \hat{c}_N = (B^{-1}N)^t \hat{c}_B \\ &= \begin{pmatrix} 0 & -1/10 & 1/10 \\ 1 & 7/10 & -7/10 \end{pmatrix} \begin{pmatrix} 20 \\ 0 \\ 30 \end{pmatrix} = \begin{pmatrix} 3 \\ -1 \end{pmatrix} \end{aligned}$$

Novas variáveis

- Com

$$B^{-1}b = \begin{pmatrix} 6000 \\ 2600 \\ 1400 \end{pmatrix}$$

- Novo valor da função objetivo

$$\hat{z}^* = \hat{c}_B^t B^{-1} b = \begin{pmatrix} 20 & 0 & 30 \end{pmatrix} \begin{pmatrix} 6000 \\ 2600 \\ 1400 \end{pmatrix} = 162000$$

Exemplo: Novo dicionário

- Novo sistema primal viável, mas não ótimo:

$$\begin{array}{llll} z & = 162000 & -3w_1 & +w_2 \\ \hline p & = 6000 & & -w_2 \\ w_3 & = 2600 & +1/10w_1 & -7/10\mathbf{w_2} \\ c & = 1400 & -1/10w_1 & +7/10w_2 \end{array}$$

- Depois um pivot: Sistema ótimo.

$$\begin{array}{llll} z & = 165714\,2/7 & -20/7w_1 & -10/7w_3 \\ \hline p & = 2285\,5/7 & -1/7w_1 & +10/7w_3 \\ w_2 & = 3714\,2/7 & +1/7w_1 & -10/7w_2 \\ c & = 4000 & & -w_3 \end{array}$$

◊

Exemplo 3.5
O que acontece se mudarmos o lucro das placas de R\$ 25 para R\$ 35 e dos canos de R\$ 30 para R\$ 10?

Exemplo: Placas e canos com lucro R\$ 35 e R\$ 10

- Novos vetores

$$\hat{c} = \begin{pmatrix} 35 \\ 0 \\ 10 \\ 0 \\ 0 \end{pmatrix}; \hat{c}_B = \begin{pmatrix} 35 \\ 0 \\ 10 \end{pmatrix}; \hat{c}_N = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

- Aumento

$$\hat{y}_N^* = ((B^{-1}N)^t c_B - c_N) = \begin{pmatrix} 0 & -1/10 & 1/10 \\ 1 & 7/10 & -7/10 \end{pmatrix} \begin{pmatrix} 35 \\ 0 \\ 10 \end{pmatrix} = \begin{pmatrix} 1 \\ 28 \end{pmatrix} =$$

Novas variáveis e novo dicionário

- Novo valor da função objetivo

$$\hat{z}^* = \hat{c}_B^t B^{-1} b = \hat{c}_B^t x_B^* = \begin{pmatrix} 35 & 0 & 10 \end{pmatrix} \begin{pmatrix} 6000 \\ 2600 \\ 1400 \end{pmatrix} = 224000$$

- O novo sistema primal viável é

$$
\begin{array}{lllr}
z & = 224000 & -1w_1 & -28w_2 \\
\hline
p & = 6000 & & -w_2 \\
w_3 & = 2600 & +1/10w_1 & -7/10w_2 \\
c & = 1400 & -1/10w_1 & +7/10w_2
\end{array}
$$

- O sistema é ótimo.

$\Diamond$

4 Tópicos

- Formalização de problemas, modelagem de problemas.
- Aplicações.
- Técnicas: Maximização do mínimo ou máximo de um conjunto finito de valores.

4.1 Função objetivo linear por segmentos

A função objetivo de um programa linear deve ser linear. Caso o problema possui uma função objetivo de outra classe (por exemplo quadrática), em geral o método Simplex não se aplica mais e a solução se torna mais difícil. Porém, nessa capítulo vamos conhecer exeções importantes: *Funções lineares por segmentos* (ingl. piecewise linear functions) e *funções racionais*.
Uma função f linear por segmentos com n segmentos é definido por

$$f(x) = f_i(x) \qquad \text{para } v_{i-1} \leq x \leq v_i\text{n} \tag{4.1}$$

usando n funções lineares $f_1, \ldots, f_n$ e $n-1$ pontos de separação $v_1, \ldots, v_{n-1}$ tal que $f_i(v_1) = f_{i+1}(v_i)$ para $1 \leq i < n$ (com pontos auxiliares $v_0 = -\infty$ e $v_n = \infty$); veja figura 4.1.
Para avaliar $f(x)$ temos que saber qual o segmento "ativo" (o segmento que contém o valor atual). Uma idéia expressar $f(x)$ como função linear é separar a variável x em n variáveis x_i positivas (exceto x_1) que representam a contribuição do segmento i ao valor total do x, tal que

$$x = \sum_{1 \leq i \leq n} x_i$$

e otimizar a função objetivo

$$f_1(x_1) + (f_2(v_1 + x_2) - f_2(v_1)) + \cdots + (f_n(v_{n-1} + x_n) - f_n(v_{n-1})). \tag{4.2}$$

com restrições

$$\begin{aligned} x_1 &\leq v_1 \\ 0 \leq x_i &\leq v_i - v_{i-1} \qquad\qquad 2 \leq i \leq n \end{aligned}$$

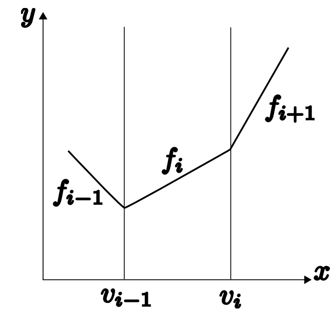


Figura 4.1: Representação de funções lineares por segmentos.

O problema com essa abordagem é garantir, que x é representado "da esquerda à direita", i.e. antes de atribuir um valor à variável x_i, todas variáveis x_j com $j < i$ devem receber o seu valor máximo. O exemplo seguinte mostra, que isso nem sempre é o caso.

Exemplo 4.1
Vamos supor que queremos maximizar a função $f(x) = |x|$ no intervalo $-1 \leq x \leq 1$, i.e. resolver o problema de otimização

$$\begin{aligned} \textbf{maximiza} \quad & |x| \\ \textbf{sujeito a} \quad & -1 \leq x \leq 1 \end{aligned}$$

O ponto $v_1 = 0$ e $f_1(x) = -x$ e $f_2(x) = x$ definam os segmentos lineares dessa função conforme equação 4.1.
A função objetivo linear conforme equação 4.2 é

$$f_1(x_1) + f_2(v_1 + x_2) - f_2(v_1) = -x_1 + (x_2 - 1)$$

e o sistema completa correspondente com variáveis x_1 e x_2 seria

$$\begin{aligned} \textbf{maximiza} \quad & -x_1 + x_2 - 1 & (4.3) \\ \textbf{sujeito a} \quad & x_1 \leq 0 & (4.4) \\ & -1 \leq x_1 + x_2 \leq 1 & (4.5) \\ & x_2 \geq 0 & (4.6) \end{aligned}$$

As soluções $-x_1 = x_2 = c$ com $c \to \infty$ mostram, que esse sistema é ilimitado.
$\Diamond$

O exemplo mostrou, que a nossa abordagem não funciona no caso geral.
TBD

- Motivate, that for concave functions, the maximization would be correct, and, similarly, for convex function the minimization would work.
- Discuss why solutions of the type $x_{i+1} \leq Mx_i$ would not work in general. How to choose M?
- Give a pointer to integer programming, which allows to handle arbitrary piecewise linear functions.

4.2 Eliminação de Fourier-Motzkin

A eliminação de Fourier-Motzkin foi a primeira técnica para resolver uma sistema de desigualdades. Fourier inventou o método em 1826. Ele foi redescoberta pelo Motzkin em 1936. O objetivo é achar um soluções do sistema

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &\leq b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &\leq b_2 \\ &\cdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &\leq b_m \end{aligned}$$

(brevemente $Ax \leq b$ em notação matricial). Começamos com um exemplo. Dado

$$\begin{aligned} 2x_1 - 3x_2 &\leq 5 \\ 4x_1 + 8x_2 &\leq 20 \end{aligned}$$

nos queremos eliminar a variável x_2. Reescrevendo as desigualdades, obtemos

$$\begin{aligned} 3x_2 \geq 2x_1 - 5 &\iff x_2 \geq 2/3x_1 - 5/3 \\ 8x_2 \leq 20 - 4x_1 &\iff x_2 \leq 5/2 - 1/2x_1 \end{aligned}$$

Combinado as duas desigualdades, temos

$$2/3x_1 - 5/3 \leq 5/2 - 1/2x_1 \iff x_1 \leq 20/7$$

A partir dessa desigualdade podemos concluir, que existem soluções viáveis. Para obter uma solução concreta, podemos escolher um valor arbitrário de x_1, por exemplo $x_1 = 2$. Substituindo de novo nas primeiras equações obtemos

$$-1/3 \leq x_2 \leq 3/2$$

e podemos escolher, por exemplo, o valor $x_2 = 1$.
Em geral, a idéia é repetidamente eliminar uma variável com essa técnica até chegar numa única variável. Se as equações restantes são viáveis, o sistema original também é, e com o método de substituição podemos produzir soluções. Particularmente, para eliminar a variável x_k separamos os coeficientes a_{ik} em $L = \{i|a_{ik} < 0\}$ e $U = \{i|a_{ik} > 0\}$. Cada equação em L gera uma cota inferior para x_k e cada equação em U uma cota superior da forma

$$\begin{aligned} a_{ik}^{-1}\left(\sum_{j\neq k} a_{ij}x_j - b_i\right) &\leq x_k \qquad \text{para } i \in L \\ x_k &\leq a_{ik}^{-1}\left(b_i - \sum_{j\neq k} a_{ij}x_j\right) \qquad \text{para } i \in U \end{aligned}$$

Em seguida, podemos eliminar a variável x_k combinando cada par de cota inferior e cota superior, e ainda copiando as desigualdades que não contém x_k para o novo sistema:

$$a_{ik}^{-1}\left(\sum_{j\neq k} a_{ij}x_j - b_i\right) \leq a_{i'k}^{-1}\left(b'_i - \sum_{j\neq k} a_{i'j}x_j\right) \forall i \in L, i' \in U$$

(Observe que no caso $L = \emptyset$ ou $U = \emptyset$ não serão geradas novas desigualdades.) No pior caso nos temos $|L| = |U| = m/2$ e portanto $(m/2)^2$ equações depois uma eliminação de uma variável. Logo, em cada iteração o número de equações pode ser o quadrado do número anterior. Eliminando $n-1$ variáveis, o algoritmo possui uma complexidade pessimista de $O(m^{2^n})$.

4.3 Métodos de pontos interiores

- TBD: Karmarkar.

4.4 Relaxaçao Lagrangeana

Queremos resolver um problema de otimização na forma geral

$$\begin{aligned} \textbf{maximiza} \quad & f(x) \\ \textbf{sujeito a} \quad & g(x) = 0 \\ & x \in \mathbb{R}^n \end{aligned}$$

com funções arbitrárias $f(x)$ e $g(x)$. Uma abordagem para obter a solução ótima seria resolver a restrições $g(x) = 0$, i.e. $V = \{x \mid g(x) = 0\}$ e depois maximizar sobre $f(x)$ para $x \in V$.

Exemplo 4.2

$$\begin{array}{rl} \textbf{maximiza} & 2x + y \\ \textbf{sujeito a} & x^2 + y^2 = 2 \\ & x, y\mathbb{R} \end{array}$$

Nesse exemplo podemos usar a restrição para eliminar $y = \pm\sqrt{2 - x^2}$ na função objetivo e depois maximizar $h(x) = 2x \pm \sqrt{2 - x^2}$ para $x \in \mathbb{R}$. A derivada é $h'(x) = 2 \pm x(2 - x^2)^{-1/2}$ e a condição $h'(x) = 0$ leva a solução $x = \sqrt{8/5}$ e $y = \sqrt{2/5}$.
TBD: Gráfico simples. ♢

TBD: I tried to complicate the example, using

$$\begin{array}{rl} \textbf{maximiza} & -x - y \\ \textbf{sujeito a} & x^2 + y^2 = e^{cx+y} \\ & x, y \in \mathbb{R} \end{array} \qquad \begin{array}{r} (4.7) \\ (4.8) \\ (4.9) \end{array}$$

but the Lagrangian approach makes the example not _much_more solvable.
TBD: Lagrangian approach.

Abordagem Lagrangeana Supomos que x^* é uma solução ótima que satisfaz $g(x^*) = 0$. Logo, o normal das restrições $g(x)$ é na direção do gradiente da função objetivo $f(x)$, senão a função objetivo não é um máximo local. Isso permite obter a solução ótimo como solução do sistema

$$\begin{aligned} \nabla f(x) &= \lambda \nabla g(x) \\ g(x) &= 0 \\ \lambda &\geq 0 \end{aligned}$$

de $n + 1$ equações e inconhecidas. O *multiplicador Lagrangeana* λ representa a fração entre o tamanho de ∇f e ∇g.
TBD: Gráfico.
TBD: Isnt this just a necessary condition for a local minimum? Wheres the optimization?

Para escrever o sistema mais compacto, vamos introduzir a *função Lagrangeana*

$$L(x, \lambda) = f(x) - \lambda g(x).$$

TBD: Check the sign, Usually (see below with multiple constraints), its positive.
TBD: The penalty interpretation.
Com isso obtemos a condição simples

$$\nabla L(x, \lambda) = 0; \qquad \lambda \geq 0$$

TBD: Exemplo.
TBD: Talk about the advantages (1) Function g(x) must not be inverted! (2) Instead of optimizing a function in n vars with k constraints, after determining the λ's, we can optimize and unconstrained function in $n + k$ vars. Give examples, which show this!

Multiplas restrições No caso geral temos um sistema com k restrições

$$\begin{array}{rll} \textbf{optimiza} & f(x) & \\ \textbf{sujeito a} & g_i(x) = 0 & 1 \leq i \leq k \\ & x \in \mathbb{R}^n & \end{array}$$

e a função Lagrageana

$$L(x, \lambda) = f(x) + \sum_{1 \leq i \leq k} \lambda_i g_i(x) \tag{4.10}$$

com multiplicadores Lagrangeanos $\lambda = (\lambda_1, \ldots, \lambda_k) \in \mathbb{R}^k$.

Aplicação à PL Dado o sistema

$$\begin{array}{rl} \textbf{maximiza} & c^t x \\ \textbf{sujeito a} & Ax = b \\ & x \geq 0 \end{array}$$

a função Lagrangeana é

$$L(x, \lambda) = c^t x + \sum_{1 \leq i \leq k} \lambda_i (a_i x - b_i)$$

e obtemos as condições

$$
\begin{aligned}
c_j + \sum_{1\le i\le k} \lambda_i a_{ij} &= 0 && 1 \le j \le n \\
a_j x - b_j &= 0 && 1 \le j \le k
\end{aligned}
$$

ou equivalente

$$
\begin{aligned}
\lambda^t A &= -c^t \\
Ax &= b
\end{aligned}
$$

TBD: Relate it better to the dual of the LP. For now, we only know that x is a primal and y is a dual solution.

5 Exercícios

(Soluções a partir da página 181.)

Exercício 5.1
Na definição da programação linear permitimos restrições lineares da forma

$$a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n \bowtie_i b_i$$

com $\bowtie_i \in \{\leq, =, \geq\}$. Porque não permitimos $\bowtie_i \in \{<, >\}$ também? Discute.

Exercício 5.2
Procura a tabela nutricional de algum restaurante e resolve o problema da dieta (exemplo 1.2).

Exercício 5.3
Um investidor pode vender ações de suas duas empresas na bolsa de valores, mas está sujeito a um limite de 10.000 operações diárias (vendas por dia). Na cotação atual, as ações da empresa A valorizaram-se 10% e agora cada uma vale R$ 22. Já a empresa B teve valorização de 2% e cada ação vale R$ 51. Sabendo-se que o investidor possui 6.000 ações da Empresa A e 7.000 da empresa B, maximize seu lucro na BOVESPA e diga qual o lucro obtido.

Exercício 5.4
Dona Maria adora ver seus netinhos Marcos, Renato e Vinicius bem alimentados. Sempre na hora de cozinhar ela leva em conta o quanto eles gostam de cada prato para fazê-los comer o máximo possível. Marcos gosta da lasanha e comeria 3 pratos dela após um prato de sopa; Renato prefere lanches, e comeria 5 hambúrgueres, ignorando a sopa; Vinicius gosta muita da massa a bolonhesa, e comeria 2 pratos após tomar dois pratos de sopa. Para fazer a sopa, são necessários entre outros ingredientes, 70 gramas de queijo por prato e 30 gramas de carne. Para cada prato de lasanha, 200 gramas de queijo, e 100 gramas de carne. Para cada hambúrguer são necessários 100 gramas de carne, e 100 gramas de queijo. Para cada prato de massa a bolonhesa so necessários 100 gramas de carne e 30 gramas de queijo (ralado para por sobre a massa). Seus netos vieram visitá-la de surpresa, e tendo ela somente 800 gramas de carne e 1000 gramas de queijo em casa, como ela poderia fazê-los comer o maior número de pratos, garantindo que cada um deles comerá pelo menos dois pratos, e usando somente os ingredientes que ela possui?

Exercício 5.5
Uma cidade foi invadida por pombas, e o prefeito obteve licença para eliminar 35.000 pombas durante apenas um dia. Ele pode usar duas estratégias: ou contratar adultos que poderão portar uma arma de matar pombas, ou crianças com "bodoques". O prefeito tem orçamento para pagar no máximo 340 adultos, sendo que 200 se inscreveram para a missão, enquanto que 7000 crianças se inscreveram. Estima-se que um adulto consiga matar 40 pombas por dia, enquanto que uma criança apenas 5. Cada adulto receberá um valor de R$30 pelo dia de trabalho, enquanto que as crianças receberão R$10. Quantos adultos e quantas crianças devem ser contratados de forma que o serviço seja feito, e seja gasto o menor custo possível com a missão?

Exercício 5.6 ([4])
Formalize como problema de otimização linear e resolve graficamente.
Uma empresa de aço produz placas ou canos de ferro. As taxas de produção são 200t/h para placas e 140t/h para canos. O lucro desses produtos e 25$/t para placas e 30$/t para canos. Considerando a demanda atual, os limites de produção são 6000t de placas e 4000t de canos. Na semana atual são 40h de tempo de produção disponível. Quantas toneladas de placas e canos devem ser produzidas para maximizar o lucro?

Exercício 5.7 ([4])
Formalize como problema de otimização linear.
Uma pequena empresa aérea oferece um vôo de Pelotas, com escala em Porto Alegre para Torres. Logo tem três tipos de clientes que voam Pelotas–Porto Alegre, Pelotas–Torres e Porto Alegre–Torres. A linha também oferece três tipos de bilhetes:

- Tipo A: bilhete regular.
- Tipo B: sem cancelamento.
- Tipo C: sem cancelamento, pagamento três semanas antes de viajar.

Os preços (em R$) dos bilhetes são os seguintes

	Pelotas–Porto Alegre	Porto Alegre–Torres	Pelotas–Torres
A	600	320	720
B	440	260	560
C	200	160	280

Baseado em experiência com esse vôo, o marketing tem a seguinte predição de passageiros:

	Pelotas–Porto Alegre	Porto Alegre–Torres	Pelotas–Torres
A	4	8	3
B	8	13	10
C	22	20	18

O objetivo da empresa e determinar o número ótimo de bilhetes para vender de cada tipo, respeitando um limite de 30 passageiros em cada vôo e o limite dos passageiros previstos em cada categoria, que maximiza o lucro.

Exercício 5.8
Escreve em forma normal.

$$\begin{aligned} \textbf{minimiza} \quad & z = -5x_1 - 5x_2 - 5x_3 \\ \textbf{sujeito a} \quad & -6x_1 - 2x_2 - 9x_3 \le 0 \\ & -9x_1 - 3x_2 + 3x_3 = 3 \\ & x_j \ge 0 \end{aligned}$$

$$\begin{aligned} \textbf{maximiza} \quad & z = -6x_1 - 2x_2 - 6x_3 + 4x_4 + 4x_5 \\ \textbf{sujeito a} \quad & -3x_1 - 8x_2 - 6x_3 - 7x_4 - 5x_5 = 3 \\ & 5x_1 - 7x_2 + 7x_3 + 7x_4 - 6x_5 \le 6 \\ & 1x_1 - 9x_2 + 5x_3 + 7x_4 - 10x_5 = -6 \\ & x_j \ge 0 \end{aligned}$$

$$\begin{aligned} \textbf{maximiza} \quad & z = 7x_1 + 4x_2 + 8x_3 + 7x_4 - 9x_5 \\ \textbf{sujeito a} \quad & -4x_1 - 1x_2 - 7x_3 - 8x_4 + 6x_5 = -2 \\ & x_1 + 4x_2 + 2x_3 + 2x_4 - 7x_5 \ge -7 \\ & -8x_1 + 2x_2 + 8x_3 - 6x_4 - 7x_5 = -7 \\ & x_j \ge 0 \end{aligned}$$

$$\begin{aligned} \textbf{minimiza} \quad & z = -6x_1 + 5x_2 + 8x_3 + 7x_4 - 8x_5 \\ \textbf{sujeito a} \quad & -5x_1 - 2x_2 + x_3 - 9x_4 - 7x_5 = 9 \\ & 7x_1 + 7x_2 + 5x_3 - 3x_4 + x_5 = -8 \\ & -5x_1 - 3x_2 - 5x_3 + 9x_4 + 8x_5 \le 0 \\ & x_j \ge 0 \end{aligned}$$

Exercício 5.9 ([3])
Resolve com o método Simplex.

$$\begin{aligned} \textbf{maximiza} \quad & z = 3x_1 + 5x_2 \\ \textbf{sujeito a} \quad & x_1 \le 4 \\ & x_2 \le 6 \\ & 3x_1 + 2x_2 \le 18 \\ & x_j \ge 0 \end{aligned}$$

Exercício 5.10
Resolve o exercício 5.6 usando o método Simplex.

Exercício 5.11
Prova que

$$\frac{2^{2n}}{2n} \le \binom{2n}{n} \le 2^{2n}.$$

Exercício 5.12
Resolve o sistema degenerado

$$\begin{array}{llllll} z & = & 10x_1 & -57x_2 & -9x_3 & -24x_4 \\ \hline w_1 & = & -1/2x_1 & +11/2x_2 & +5/2x_3 & -9x_4 \\ w_2 & = & -1/2x_1 & +3/2x_2 & +1/2x_3 & -x_4 \\ w_3 & = 1 & -x_1 & & & \end{array}$$

usando o método lexicográfico e o regra de Bland.

Exercício 5.13
Dado o problema de otimização

$$\begin{aligned} \textbf{maximiza} \quad & x_1 + x_2 \\ \textbf{sujeito a} \quad & ax_1 + bx_2 \le 1 \\ & x_1, x_2 \ge 0 \end{aligned}$$

determine condições suficientes e necessárias ao respeita de a e b tal que

1. existe ao menos uma solução ótima,
2. existe exatamente uma solução ótima,
3. existe nenhuma solução ótima,

4. o sistema é ilimitado.

ou demonstre que o caso não é possível.

Exercício 5.14
Sabe-se que o dicionário ótimo do problema

$$\begin{array}{rl} \textbf{maximiza} & z = 3x_1 + x_2 \\ \textbf{sujeito a} & 2x_1 + 3x_2 \leq 5 \\ & x_1 - x_2 \leq 1 \\ & x_1, x_2 \geq 0 \end{array}$$

é

$$\begin{array}{llll} z^* & = 31 & -11w_2 & -4w_1 \\ \hline x_2 & = 7 & -2w_2 & -w_1 \\ x_1 & = 8 & -3w_2 & -w_1 \end{array}$$

- Se a função objetivo passar a $z = x_1 + 2x_2$, a solução continua ótima? No caso de resposta negativa, determine a nova solução ótima.
- Se a função objetivo passar a $z = x_1 - x_2$, a solução continua ótima? No caso de resposta negativa, determine a nova solução ótima.
- Se a função objetivo passar a $z = 2x_1 - 2x_2$, a solução continua ótima?No caso de resposta negativa, determine a nova solução ótima.
- Formular o dual e obter a solução dual ótima.