Programação Linear - método simplex: situações particulares

Investigação Operacional

J.M. Valério de Carvalho
vc@dps.uminho.pt

Departamento de Produção e Sistemas
Escola de Engenharia, Universidade do Minho

6 de outubro de 2017

Prog. Linear - método simplex: situações particulares

antes

- O algoritmo Simplex foi aplicado para resolver um problema de programação linear.

Guião

- Há situações particulares em que é necessário detalhar as regras e estabelecer decisões e operações suplementares:
  - quando há degenerescência (várias bases diferentes correspondem à mesma solução básica);
  - quando o domínio é ilimitado;
  - quando não existe um vértice inicial admissível.

depois

- Analisaremos a implementação do método simplex usando matrizes.

Situações particulares do método simplex:

- Selecção de um vértice admissível inicial
  - Se não existir, **problema é impossível**.
- Repetir
  - Selecção da coluna pivô:
    - Coeficiente mais negativo na linha da função objectivo
    - (em caso de empate, escolha arbitrária)
    - Se não existir coef.<0, solução óptima.
  - Selecção da linha pivô:
    - Menor razão (lado direito/coluna pivô) positiva (coef.col.>0)
    - (em caso de empate, há **degenerescência)**
    - Se não existir coef.col.>0, **solução óptima é ilimitada**.
  - Fazer eliminação de Gauss
- Enquanto (solução não for óptima)

- O que é um algoritmo?
- O algoritmo simplex converge?

Ver Mais

Conteúdo

1. Degenerescência

- O que é a degenerescência?
- Como identificá-la no quadro simplex?
- Como escolher a linha pivô quando há empate na menor razão positiva ?

Degenerescência: o que é?

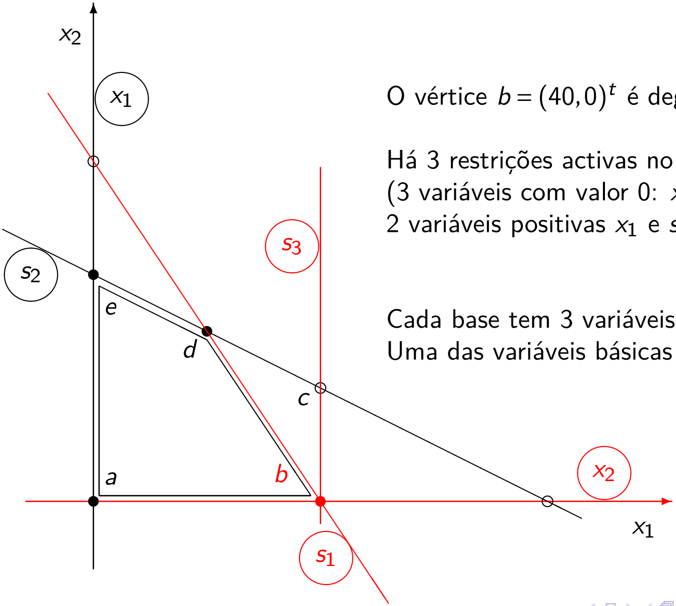
Vértice degenerado: geometria

- Normalmente, o vértice é definido pela intersecção de $(n-m)$ hiperplanos (*i.e.*, há $(n-m)$ restrições activas).
- Um vértice é *degenerado* se o número de hiperplanos for maior.

Vértice degenerado: várias bases, a mesma solução básica

- Uma *base* é um conjunto de variáveis básicas (de vectores linearmente independentes).
- Resolvendo o sistema de equações em ordem às variáveis básicas (da base), obtém-se uma solução básica (vértice).
- Pode haver várias bases (quadros simplex) cuja solução corresponda ao mesmo vértice (solução básica), que é *degenerado*.

Exemplo



O vértice $b = (40, 0)^t$ é degenerado.

Há 3 restrições activas no vértice b (3 variáveis com valor 0: x_2, s_1 e s_3) e 2 variáveis positivas x_1 e s_2.

Cada base tem 3 variáveis básicas.
Uma das variáveis básicas é nula.

Ver mais

Três bases diferentes, o mesmo vértice *(solução básica)*

	z	x_1	x_2	s_1	s_2	s_3	
s_1	0	0	2	1	0	-3	0
s_2	0	0	2	0	1	-1	40
x_1	0	1	0	0	0	1	40
z	1	0	-10	0	0	12	480

	z	x_1	x_2	s_1	s_2	s_3	
x_2	0	0	1	0.5	0	-1.5	0
s_2	0	0	0	-1	1	2	40
x_1	0	1	0	0	0	1	40
z	1	0	0	5	0	-3	480

Um quadro simplex corresponde a um vértice degenerado se houver uma ou mais variáveis básicas com valor 0.

	z	x_1	x_2	s_1	s_2	s_3	
s_2	0	0	4/3	-1/3	1	0	40
s_3	0	0	-2/3	-1/3	0	1	0
x_1	0	1	2/3	1/3	0	0	40
z	1	0	-2	4	0	0	480

Solução básica é sempre $(x_1, x_2, s_1, s_2, s_3)^t = (40, 0, 0, 40, 0)^t$.

Escolha da linha pivô quando há empate

	z	x_1	x_2	s_1	s_2	s_3		
s_1	0	$\boxed{3}$	2	1	0	0	120	$120/3 = 40$
s_2	0	1	2	0	1	0	80	$80/1 = 80$
s_3	0	1	0	0	0	1	40	$40/1 = 40$
z	1	-12	-10	0	0	0	0	

- Há empate na menor razão positiva $= 40$ (linhas de s_1 e de s_3).

Desempate:

- perturbar o lado direito, adicionando ϵ, e
- calcular novamente a menor razão positiva.

- Exemplo:
  - Linha de s_1 : $(120 + \epsilon)/3 = 40 + \epsilon/3$
  - Linha de s_3 : $(40 + \epsilon)/1 = 40 + \epsilon$
- Linha pivô correcta: a de s_1 (menor razão positiva)

Esta escolha tipicamente reduz o número de iterações.

Exemplo: resolução

	z	x_1	x_2	s_1	s_2	s_3	
s_1	0	$\boxed{3}$	2	1	0	0	120
s_2	0	1	2	0	1	0	80
s_3	0	1	0	0	0	1	40
z	1	-12	-10	0	0	0	0

	z	x_1	x_2	s_1	s_2	s_3	
x_1	0	1	2/3	1/3	0	0	40
s_2	0	0	$\boxed{4/3}$	-1/3	1	0	40
s_3	0	0	-2/3	-1/3	0	1	0
z	1	0	-2	4	0	0	480

	z	x_1	x_2	s_1	s_2	s_3	
x_1	0	1	0	0.5	-0.5	0	20
x_2	0	0	1	-0.25	0.75	0	30
s_3	0	0	0	-0.5	0.5	1	20
z	1	0	0	3.5	1.5	0	540

Solução óptima $(x_1, x_2, s_1, s_2, s_3)^t = (20, 30, 0, 0, 20)^t$.

2. Domínios ilimitados e soluções óptimas ilimitadas

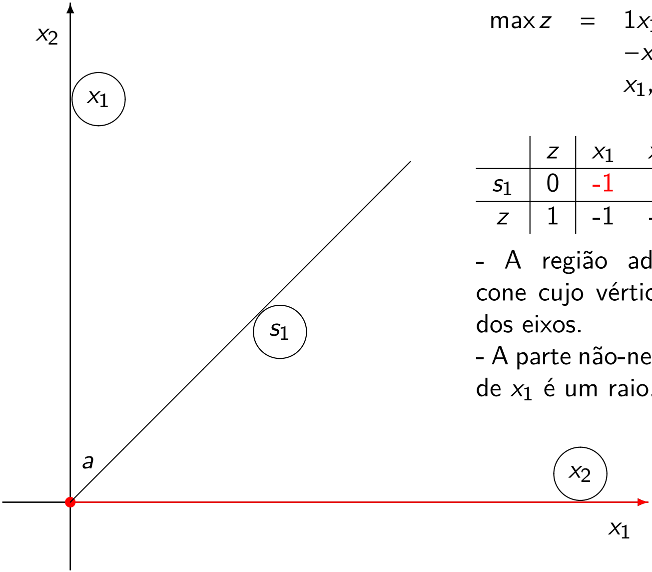
- Quando é que o domínio é ilimitado?
- Como identificar no quadro simplex?
- Se o domínio for ilimitado, a solução óptima é também ilimitada?
- Quando é que a solução óptima é ilimitada?

Domínio ilimitado (aberto)

- O domínio é ilimitado quando se pode caminhar ao longo de um raio ($\equiv$ semi-recta) permanecendo no domínio admissível.

- raio:= conjunto de pontos $\{x : x = v + \theta.d,\ \theta \in \mathrm{IR}_+\}$, sendo $v \in \mathrm{IR}^n$ um vértice e $d \in \mathrm{IR}^n$ uma direcção (um vector não-nulo).

Exemplo: raio $\{(0,0)^t + \theta(1,0)^t, \theta \geq 0\}$ (eixo de x_1)



$$\begin{aligned} \max z &= 1x_1 + 1x_2 \\ & -x_1 + x_2 \leq 0 \\ & x_1, x_2 \geq 0 \end{aligned}$$

	z	x_1	x_2	s_1	
s_1	0	-1	1	1	0
z	1	-1	-1	0	0

- A região admissível é o cone cujo vértice é a origem dos eixos.
- A parte não-negativa do eixo de x_1 é um raio.

Domínio ilimitado: como identificar no quadro simplex?

Quadro simplex: como identificar um raio?

- Há uma coluna de uma variável não-básica em que os coeficientes das restrições são todos ≤ 0.

- Exemplo:

	z	x_1	x_2	s_1	
s_1	0	-1	1	1	0
z	1	-1	-1	0	0

Ao longo de um raio, todos os pontos são admissíveis, porque:

- há uma **única** variável não-básica que aumenta de valor,
- todas as vars básicas aumentam (coef.<0) ou mantêm o valor (coef.=0),
- sendo portanto $x, s \geq 0$.

Domínio ilimitado: solução óptima ilimitada

- O valor da solução óptima é ilimitado quando, ao caminhar ao longo de um raio, o valor da função objectivo aumenta (prob. de max.).

Quadro simplex: como identificar uma solução óptima ilimitada?

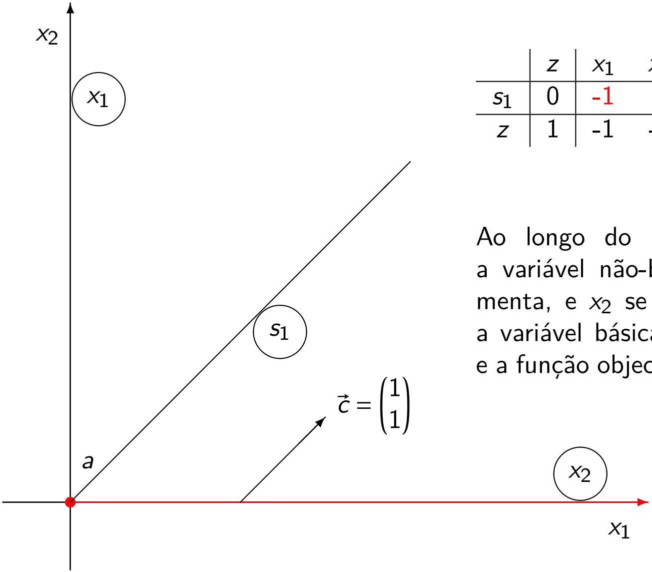
- há um raio e
- o respectivo coeficiente na linha da função objectivo é < 0.

- Exemplo:

	z	x_1	x_2	s_1	
s_1	0	-1	1	1	0
z	1	-1	-1	0	0

- Quando o coeficiente da linha da função objectivo do raio for ≥ 0, o valor da solução óptima pode não ser ilimitado (porquê?).
- Para a solução óptima ser ilimitada, o domínio deve ser ilimitado (porquê?)

Exemplo: domínio ilimitado e solução óptima ilimitada



	z	x_1	x_2	s_1	
s_1	0	-1	1	1	0
z	1	-1	-1	0	0

Ao longo do raio, quando a variável não-básica x_1 aumenta, e x_2 se mantém $=0$, a variável básica s_1 aumenta e a função objectivo também.

3. Vértice admissível inicial

- E se não houver um vértice admissível (quadro simplex) inicial?
- Exemplo: problema com restrições de $\geq$.
- O Método das 2 Fases
  - Fase I: obter um vértice admissível inicial
  - Fase II: aplicar algoritmo simplex

Um problema com restrições de $\geq$ e de minimização

$$
\begin{array}{lrcrcrcl}
\min z = & 120y_3 & + & 80y_4 & + & 30y_5 & & \\
 & 3y_3 & + & 1y_4 & + & 1y_5 & \geq & 12 \\
 & 2y_3 & + & 2y_4 & & & \geq & 10 \\
 & \multicolumn{7}{l}{y_3, y_4, y_5 \geq 0}
\end{array}
$$

- Há uma relação este problema e o que vimos anteriormente.
- Iremos explorar essa relação depois.

Transformação na forma canónica

$$
\begin{array}{lcl}
\min z & = & cx \\
 & & Ax \geq b \\
 & & x \geq 0
\end{array}
\quad \rightarrow \quad
\begin{array}{lcl}
\min z & = & cx \\
 & & Ax - u = b \\
 & & x, u \geq 0
\end{array}
$$

sendo $u \in \mathrm{IR}_+^{m\times 1}$ um vector de variáveis de folga da mesma dimensão que $b \in \mathrm{IR}^{m\times 1}$.

Transformação Inequações → Equações

- Qualquer inequação do tipo $\geq$ pode ser transformada numa equação (equivalente), introduzindo uma variável adicional, designada por *variável de folga*, com valor não-negativo.
- Exemplo:

$$\begin{array}{llll} 3x_1 + 2x_2 & \geq 120, & x_1, x_2 & \geq 0 \\ 3x_1 + 2x_2 - 1u_1 & = 120, & x_1, x_2, u_1 & \geq 0 \end{array}$$

- O número de unidades produzidas numa solução $(x_1, x_2)^t$ é igual ao valor da função linear: $3x_1 + 2x_2$.
- u_1 (variável de folga) é o número de unidades produzidas em excesso relativamente às necessárias (no exemplo, 120).
- Há autores que designam estas variáveis por *variáveis de excesso*.

Exemplo: transformação na forma canónica

Modelo original

- Variáveis de decisão: y_3, y_4, y_5.

$$
\begin{array}{lrcrcrcc}
\min z = & 120y_3 & + & 80y_4 & + & 30y_5 & & \\
 & 3y_3 & + & 1y_4 & + & 1y_5 & \geq & 12 \\
 & 2y_3 & + & 2y_4 & & & \geq & 10 \\
 & \multicolumn{7}{l}{y_3, y_4, y_5 \geq 0}
\end{array}
$$

Modelo na forma canónica (equivalente ao modelo original)

- Variáveis de decisão: y_3, y_4, y_5.
- Variáveis de excesso: y_1, y_2.

$$
\begin{array}{lrrcrcrcrl}
\min z = & & & & 120y_3 & + & 80y_4 & + & 30y_5 & \\
 & -1y_1 & & + & 3y_3 & + & 1y_4 & + & 1y_5 & = 12 \\
 & & -1y_2 & + & 2y_3 & + & 2y_4 & & & = 10 \\
 & \multicolumn{9}{l}{y_1, y_2, y_3, y_4, y_5 \geq 0}
\end{array}
$$

Não há um vértice admissível inicial, porque ...

- o lado direito é positivo e não há uma matriz identidade no quadro :

	z	y_1	y_2	y_3	y_4	y_5	
	0	-1	0	3	1	1	12
	0	0	-1	2	2	0	10

Lembrete: antes havia um vértice admissível inicial, porque:

- as restrições eram todas de $\leq$ (havia uma matriz identidade $I_{m \times m}$), e
- os coeficientes do lado direito eram todos ≥ 0.

- Quando não há um vértice admissível inicial, usa-se o:

Método das 2 Fases:

- na Fase I, resolve-se um *problema auxiliar* para tentar encontrar um vértice admissível inicial.
- Se se conseguir encontrar, na Fase II, aplica-se o algoritmo simplex; caso contrário, o problema é impossível.

Método das 2 fases: estratégia

Fase I: adicionar variáveis artificiais e minimizar a sua soma

- resolver problema auxiliar ($\mathbf{1}a$ é a soma das variáveis artificiais):

$$\begin{aligned} \min z_a &= \mathbf{1}a \\ & Ax - u + a = b \\ & x, u, a \geq 0 \end{aligned}$$

  sendo $a \in \mathrm{IR}_+^{m\times 1}, \mathbf{1} = [1, 1, \ldots, 1]$ um vector linha com m elementos.
- Se $(\min z_a = \mathbf{1}a = 0) \Rightarrow a = \tilde{0}$ (todas as variáveis artificiais $= 0$) $\Rightarrow$ há um vértice admissível que obedece às restrições originais;
- caso contrário ($\min z_a > 0$), não é possível obter uma solução que obedeça a todas as restrições originais $\Rightarrow$ problema é impossível.

Fase II: optimizar problema original

- Existe um vértice admissível inicial para o algoritmo simplex;
- optimiza-se a função objectivo (original) do problema.

Fase I: adicionar vars artificiais a_1 e a_2, e $\min z_a$

- Função objectivo da Fase I: $\min z_a \;=\; 1a_1 + 1a_2$.
- Equação da linha da função objectivo: $z_a - 1a_1 - 1a_2 = 0$

	z_a	y_1	y_2	y_3	y_4	y_5	a_1	a_2	
a_1	0	-1	0	3	1	1	1	0	12
a_2	0	0	-1	2	2	0	0	1	10
z_a	1	0	0	0	0	0	- 1	- 1	0

- O quadro não é válido: os coeficientes da linha da função objectivo debaixo da matriz identidade devem ser nulos. Validar Quadro
- O quadro seguinte é válido; vamos minimizar a função auxiliar z_a:

	z_a	y_1	y_2	y_3	y_4	y_5	a_1	a_2	
a_1	0	-1	0	$\boxed{3}$	1	1	1	0	12
a_2	0	0	-1	2	2	0	0	1	10
z_a	1	-1	-1	5	3	1	0	0	22

Fase I: iterações

	z_a	y_1	y_2	y_3	y_4	y_5	a_1	a_2	
a_1	0	-1	0	$\boxed{3}$	1	1	1	0	12
a_2	0	0	-1	2	2	0	0	1	10
z_a	1	-1	-1	5	3	1	0	0	22

	z_a	y_1	y_2	y_3	y_4	y_5	a_1	a_2	
y_3	0	-1/3	0	1	1/3	1/3	1/3	0	4
a_2	0	2/3	-1	0	$\boxed{4/3}$	-2/3	-2/3	1	2
z_a	1	2/3	-1	0	4/3	-2/3	-5/3	0	2

	z_a	y_1	y_2	y_3	y_4	y_5	a_1	a_2	
y_3	0	-1/2	1/4	1	0	1/2	1/2	-1/4	7/2
y_4	0	1/2	-3/4	0	1	-1/2	-1/2	3/4	3/2
z_a	1	0	0	0	0	0	-1	-1	0

- Solução óptima: $\min z_a = 0$.
- Foi encontrado um vértice admissível.

Fase I: conclusão

- O vértice admissível é $(y_1, y_2, y_3, y_4, y_5)^t = (0, 0, 7/2, 3/2, 0)^t$.
- Variáveis artificiais (a_1, a_2) e função objectivo auxiliar (z_a) não são necessárias na Fase II, e podem ser eliminadas.

		y_1	y_2	y_3	y_4	y_5	
y_3		-1/2	1/4	1	0	1/2	7/2
y_4		1/2	-3/4	0	1	-1/2	3/2

- Na Fase II, iremos optimizar a função objectivo original (z), partindo do vértice admissível encontrado na Fase I.

Fase II: função objectivo original

- Função objectivo da Fase II: min $z = 120y_3 + 80y_4 + 30y_5$.
- Equação da linha da função objectivo: $z - 120y_3 - 80y_4 - 30y_5 = 0$

	z	y_1	y_2	y_3	y_4	y_5	
y_3	0	-1/2	1/4	1	0	1/2	7/2
y_4	0	1/2	-3/4	0	1	-1/2	3/2
z	1	0	0	-120	-80	-30	0

- O quadro não é válido: os coeficientes da linha da função objectivo debaixo da matriz identidade devem ser nulos.

Validar Quadro

- O quadro seguinte é válido; vamos optimizar a função original z:

	z	y_1	y_2	y_3	y_4	y_5	
y_3	0	-1/2	1/4	1	0	1/2	7/2
y_4	0	1/2	-3/4	0	1	-1/2	3/2
z	1	-20	-30	0	0	-10	540

- O primeiro vértice admissível encontrado é a solução óptima
- (problema de minimização e nenhum coeficiente na linha da função objectivo é positivo).
- Isto nem sempre acontece!

Conclusão

- Há outros algoritmos para resolver problemas de programação linear, como os métodos de pontos interiores.
- O algoritmo simplex permanece competitivo, embora tenham sido identificados problemas em que os métodos de pontos interiores têm melhor desempenho.

Apêndices

1. Método do Grande M: estratégia

associar uma penalidade muito grande às vars artificiais, para conduzi-las a um valor nulo

- resolver problema auxiliar:

$$\begin{aligned} \min z_M &= cx + \mathbf{M}a \\ & Ax - u + a = b \\ & x, u, a \geq 0 \end{aligned}$$

  sendo $a \in \mathrm{IR}_+^{m\times 1}, \mathbf{M} = [M, M, \ldots, M]$ um vector linha com m elementos.

- Se M for suficientemente grande, qualquer ponto admissível é melhor do que um ponto em que uma variável artificial seja positiva.
- Se $(a = \tilde{0})$ (todas as variáveis artificiais $= 0$) $\Rightarrow \min z_M = cx^*$ e x^* é o vértice admissível óptimo, que obedece às restrições originais.
- caso contrário ($\exists a_i > 0$), não é possível obter uma solução que obedeça a todas as restrições originais $\Rightarrow$ problema é impossível.

1. Método do Grande M: desvantagens

Se o valor de M for muito grande,

- pode haver perda de informação, resultante da representação dos números em computador.
- Os coeficientes de custo são representados por reais de dupla precisão com um número finito de casas decimais.
- Exemplo:

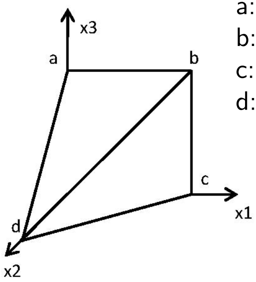
$$\begin{aligned} c_1 &= 3,1415926535897932e+00 \\ M &= 1,0000000000000000e+40 \\ M+c_1 &= 1,0000000000000000e+40 \end{aligned}$$

Se o valor de M for muito pequeno,

- pode não ser suficientemente grande para conduzir todas as variáveis artificiais a 0.

2. Degenerescência e restrições redundantes

- No exemplo, há uma *restrição redundante*, que pode ser eliminada sem alterar o domínio.
- Isso não acontece na generalidade.
- Exemplo: as restrições $x_1 + x_2 \leq 1, x_2 + x_3 \leq 1, x_1, x_2, x_3 \geq 0$ são todas necessárias:



a: $x_2 = 0$; $x_1 = 0$; $x_2 + x_3 = 1$
b: $x_2 = 0$; $x_2 + x_3 = 1$; $x_1 + x_2 = 1$
c: $x_2 = 0$; $x_3 = 0$; $x_1 + x_2 = 1$
d: $x_3 = 0$; $x_1 = 0$; $x_1 + x_2 = 1$; $x_2 + x_3 = 1$(há 4 planos)

2. Degenerescência e bases óptimas

	z'	x_1	x_2	s_1	s_2	s_3	
s_1	0	0	$\boxed{2}$	1	0	-3	0
s_2	0	0	2	0	1	-1	40
x_1	0	1	0	0	0	1	40
z'	1	0	-1	0	0	3	120

	z'	x_1	x_2	s_1	s_2	s_3	
x_2	0	0	1	0.5	0	-1.5	0
s_2	0	0	0	-1	1	2	40
x_1	0	1	0	0	0	1	40
z'	1	0	0	1/2	0	3/2	120

	z'	x_1	x_2	s_1	s_2	s_3	
s_2	0	0	4/3	-1/3	1	0	40
s_3	0	0	-2/3	-1/3	0	1	0
x_1	0	1	2/3	1/3	0	0	40
z'	1	0	1	1	0	0	120

Se $z' = 3x_1 + 1x_2$, uma das três bases da solução básica óptima não é óptima.

O quadro 1 é uma base que não é óptima: é necessário fazer um pivô degenerado para se comprovar que a solução básica (que não se altera quando se faz o pivô) é uma solução óptima.

2. Degenerescência e finitude do algoritmo simplex

- Quando não há degenerescência, o algoritmo simplex termina num número finito de iterações:

  - em cada iteração, a função objectivo melhora, e, se o óptimo for finito, o número de iterações não pode ser infinito.

- Quando há degenerescência, o algoritmo simplex pode entrar em ciclo.
- Há exemplos especialmente construídos em que a regra de seleccionar a coluna pivô com o coeficiente mais negativo pode levar a que o algoritmo entre em ciclo, percorrendo ciclicamente as diferentes bases correspondentes ao mesmo vértice.

- Para informação adicional, ver: Bland, R. "New finite pivoting rules for the simplex method". Mathematics of Operations Research 2 (2): 103 - 107, 1977. doi:10.1287/moor.2.2.103

Voltar

Algoritmo simplex de minimização

Lembrete: no algoritmo simplex de minimização:

- a coluna pivô é a coluna com o coeficiente mais positivo na linha da função objectivo,
- a solução é óptima se não existir nenhum coeficiente positivo na linha da função objectivo.

NOTA: Em alternativa a usar um algoritmo de minimização, podemos usar um algoritmo simplex de maximização para maximizar a função simétrica da função objectivo. Ver diapositivos sobre Transformações básicas.

Voltar

I - Obter quadro válido: folha de rascunho

	z_a	y_1	y_2	y_3	y_4	y_5	a_1	a_2	
a_1	0	-1	0	3	1	1	1	0	12
a_2	0	0	-1	2	2	0	0	1	10
z_a	1	0	0	0	0	0	- 1	- 1	0

- Exprimir a função objectivo z_a em função das variáveis não-básicas y_1, y_2, y_3, y_4 e y_5 usando eliminação de Gauss: somar à linha de z_a as linhas de a_1 e a_2.

	z_a	y_1	y_2	y_3	y_4	y_5	a_1	a_2	
$(+1)$*linha de z_a	1	0	0	0	0	0	-1	-1	0
$(+1)$*linha de a_1	0	-1	0	3	1	1	1	0	12
$(+1)$*linha de a_2	0	0	-1	2	2	0	0	1	10
z_a	1	-1	-1	5	3	1	0	0	22

- O quadro seguinte é válido:

Voltar

	z_a	y_1	y_2	y_3	y_4	y_5	a_1	a_2	
a_1	0	-1	0	3	1	1	1	0	12
a_2	0	0	-1	2	2	0	0	1	10
z_a	1	-1	-1	5	3	1	0	0	22

II - Obter quadro válido: folha de rascunho

	z	y_1	y_2	y_3	y_4	y_5	
y_3	0	-1/2	1/4	1	0	1/2	7/2
y_4	0	1/2	-3/4	0	1	-1/2	3/2
z	1	0	0	-120	-80	-30	0

- Exprimir a função objectivo z em função das variáveis não-básicas y_1, y_2 e y_5 usando eliminação de Gauss: somar à linha de z as linhas de y_3 e y_4 multiplicadas por constantes adequadas.

	z	y_1	y_2	y_3	y_4	y_5	
(+1)*linha de z	1	0	0	-120	-80	-30	0
(+120)*linha de y_3	0	-60	30	120	0	60	420
(+80)*linha de y_4	0	40	-60	0	80	-40	120
z	1	-20	-30	0	0	-10	540

- O quadro seguinte é válido:

Voltar

	z	y_1	y_2	y_3	y_4	y_5	
y_3	0	-1/2	1/4	1	0	1/2	7/2
y_4	0	1/2	-3/4	0	1	-1/2	3/2
z	1	-20	-30	0	0	-10	540

Fim