DUALIDADE EM PNL

PROBLEMAS: PRIMAL E DUAL

SEJA

$$(P) \qquad v = \min \ f(x) \quad \text{suj. a} \quad g(x) \leq 0, \quad x \in X \subseteq \mathbb{R}^n$$

ONDE f É UMA FUNÇÃO DE VALOR REAL

$g(x) = (g_1(x), \dots, g_m(x))$ É UMA FUNÇÃO DE $\mathbb{R}^n$ EM $\mathbb{R}^m$, E

X É UM CONJUNTO NÃO-VAZIO

OBS.: - SE NÃO EXISTIR x TAL QUE $g(x) \leq 0$ CONVENCIONA-SE $v = +\infty$

- (P) SERÁ CHAMADO <u>PRIMAL</u>
- ASSUME-SE QUE f E g SÃO CONTÍNUAS E X COMPACTO.
- SUPOR (P) SER TAL QUE $g(x) \leq 0$ O TORNAM DE DIFÍCIL SOLUÇÃO.

Função Lagrangeana: Para dado $u \geq 0$,

$$\boxed{L(u) = \min_{x \in X} \{ f(x) + u\,g(x) \} = \min_{x \in X} L(x,u)} \quad (1)$$

SEJA $\bar{x} \in X$ SOLUÇÃO PRIMAL DE (1) PARA DADO $\bar{u} \geq 0$, ISTO É $L(\bar{u}) = L(\bar{x}, \bar{u})$.

QUAL A RELAÇÃO DE $\bar{x}$ EM P ?

- ÓTIMO ?
- VIÁVEL ?

CONDIÇÕES DE OTIMALIDADE (C.O.)

UM PAR $(\bar{x}, \bar{u})$, COM $\bar{x} \in X$ E $\bar{u} \geq 0$, SATISF. CONDIÇÕES DE OTIMALIDADE PARA (P) SE

(1) $f(\bar{x}) + \bar{u}\,g(\bar{x}) = \min_{x \in X} \{ f(x) + \bar{u}\,g(x) \}$

(2) $\bar{u}\,g(\bar{x}) = 0$;

(3) $g(\bar{x}) \leq 0$.

$\therefore$ $\bar{x}$ É ÓTIMO PARA (P), POIS:

- $\bar{x}$ É VIÁVEL EM (P), $\bar{x} \in X$ E $g(\bar{x}) \leq 0$;
- SE $\tilde{x}$ É UMA SOL. VIÁVEL QUALQUER DE (P), TEMOS

$$f(\bar{x}) = f(\bar{x}) + \bar{u}\,g(\bar{x}) \leq f(\tilde{x}) + \bar{u}\,g(\tilde{x}) \leq f(\tilde{x}).$$

$\Rightarrow$ C.O. SÃO SUFICIENTES PARA QUE $\bar{x}$ SEJA ÓTIMA DE (P).

QUAL A RELAÇÃO $L(u) \; ? \; v$

PARA TODO $u \geq 0$ tem-se

$\boxed{L(u) \leq v}$, pois

- SE (P) NÃO POSSUI SOLUÇÃO VIÁVEL, $v = +\infty$
- SE $\tilde{x}$ É UMA SOLUÇÃO VIÁVEL QUALQUER DE (P)

$$L(u) \leq f(\tilde{x}) + u\, g(x) \leq f(\tilde{x})$$

pois $g(\tilde{x}) \leq 0$ E $u \geq 0$.

(como $\tilde{x}$ É QUALQUER PODE SER A ÓTIMA, $L(u) \leq v$)

ACHAR MAIOR LIMITANTE INFERIOR DE v RESOLVENDO O PROBLEMA <u>DUAL</u> DE (P)

$$(D) \qquad d = \sup \; L(u) \quad \text{suj. a} \quad u \geq 0$$

obs.: - $d \leq v$ e <u>possivelmente</u> $d < v$ onde diz-se existir um "gap" de dualidade.

- SE $(\bar{x}, \bar{u})$ SATISFAZ C.O. ENTÃO $\bar{u}$ É ÓTIMO PARA (D) E $d = v$.

Convexificação e Dualização: uma visão geométrica

⇒ O valor mínimo da f.o. do problema (P) convexificado é igual ao valor máximo da f.o. de (D)

∴ Se (P) é convexo (f, g, X convexos) e alguma condição de qualificação é satisfeita as condições de otimalidade (C.O.) serão necessárias e suficientes. (Shapiro, 1979)

Função Perturbação

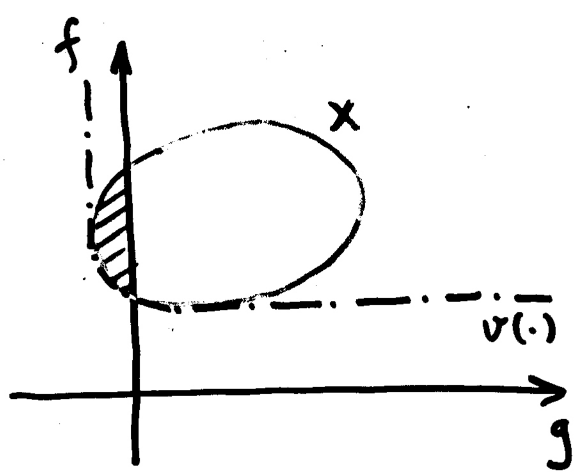
A função perturbação $v(\cdot)$ associada a (P) em $\mathbb{R}^m$

$$v(y) = \underset{x \in X}{\text{ínfimo}} \{ f(x) : \text{suj. a } g(x) \leq y \}$$

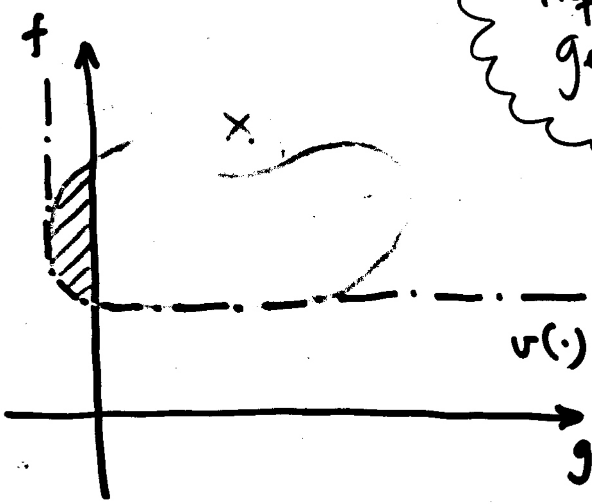
onde y é o vetor perturbação.

Pode-se mostrar que é convexa, não-crescente, semi-contínua inferior

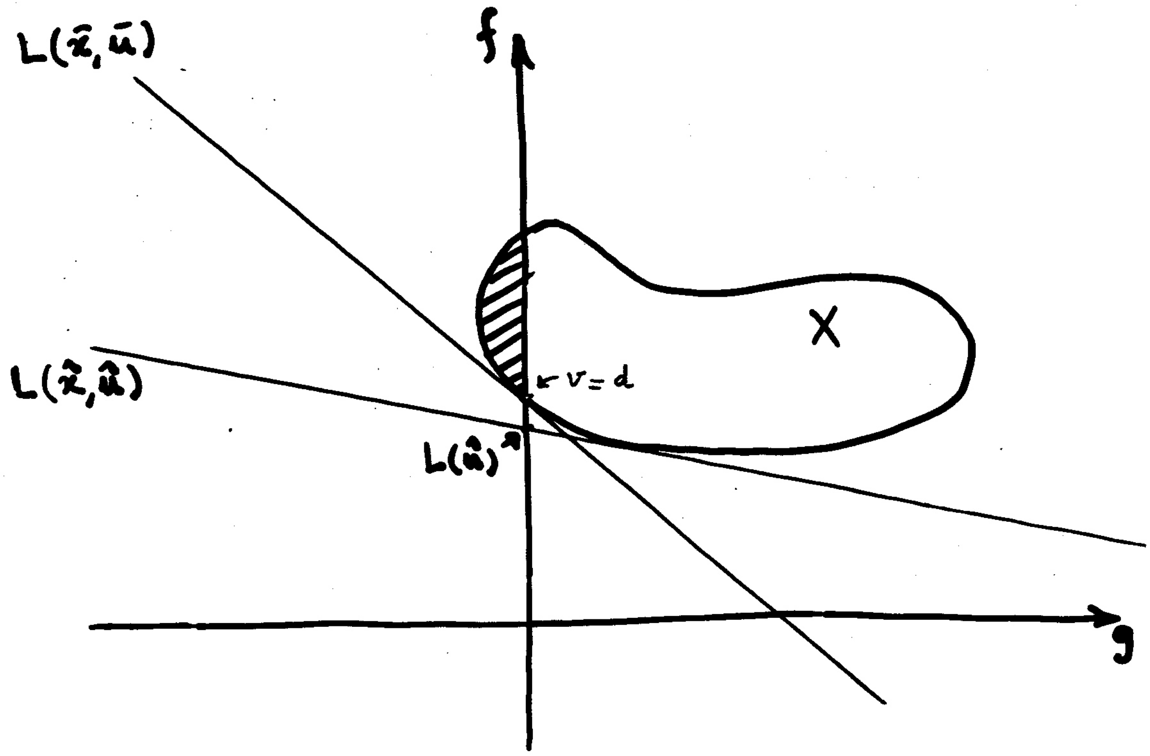
Geoffrion, 197?



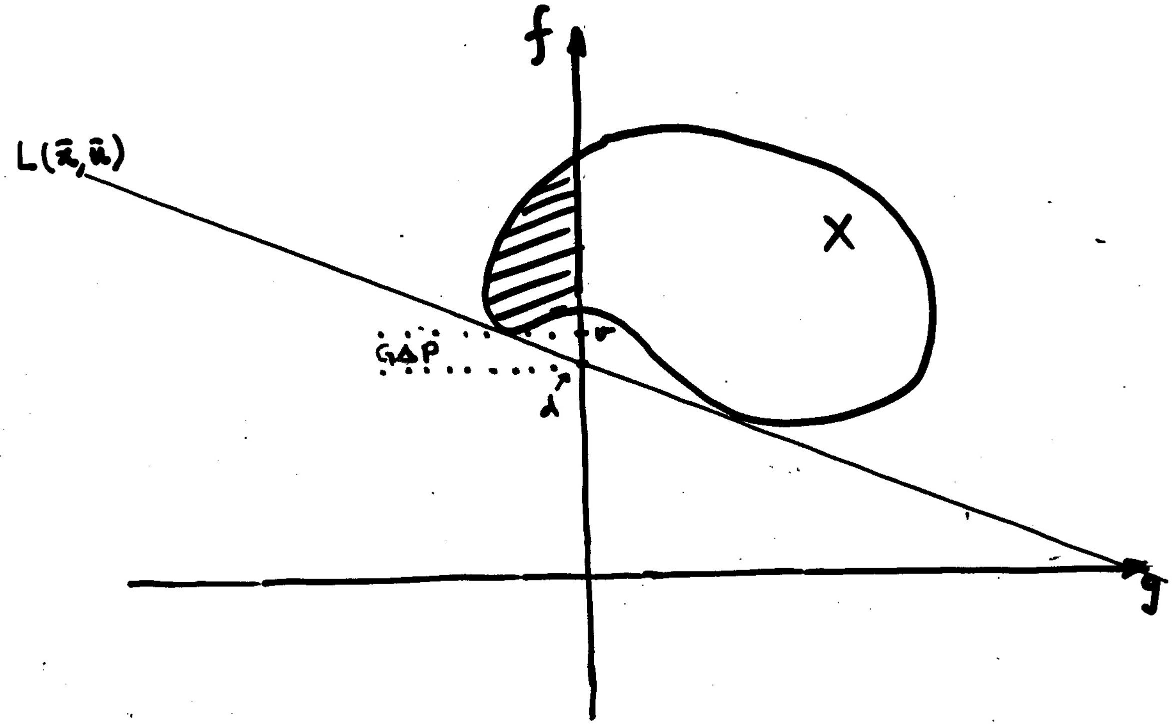
Plano (f,g) : Probl. convexo



Plano (f,g) : Probl. convexo



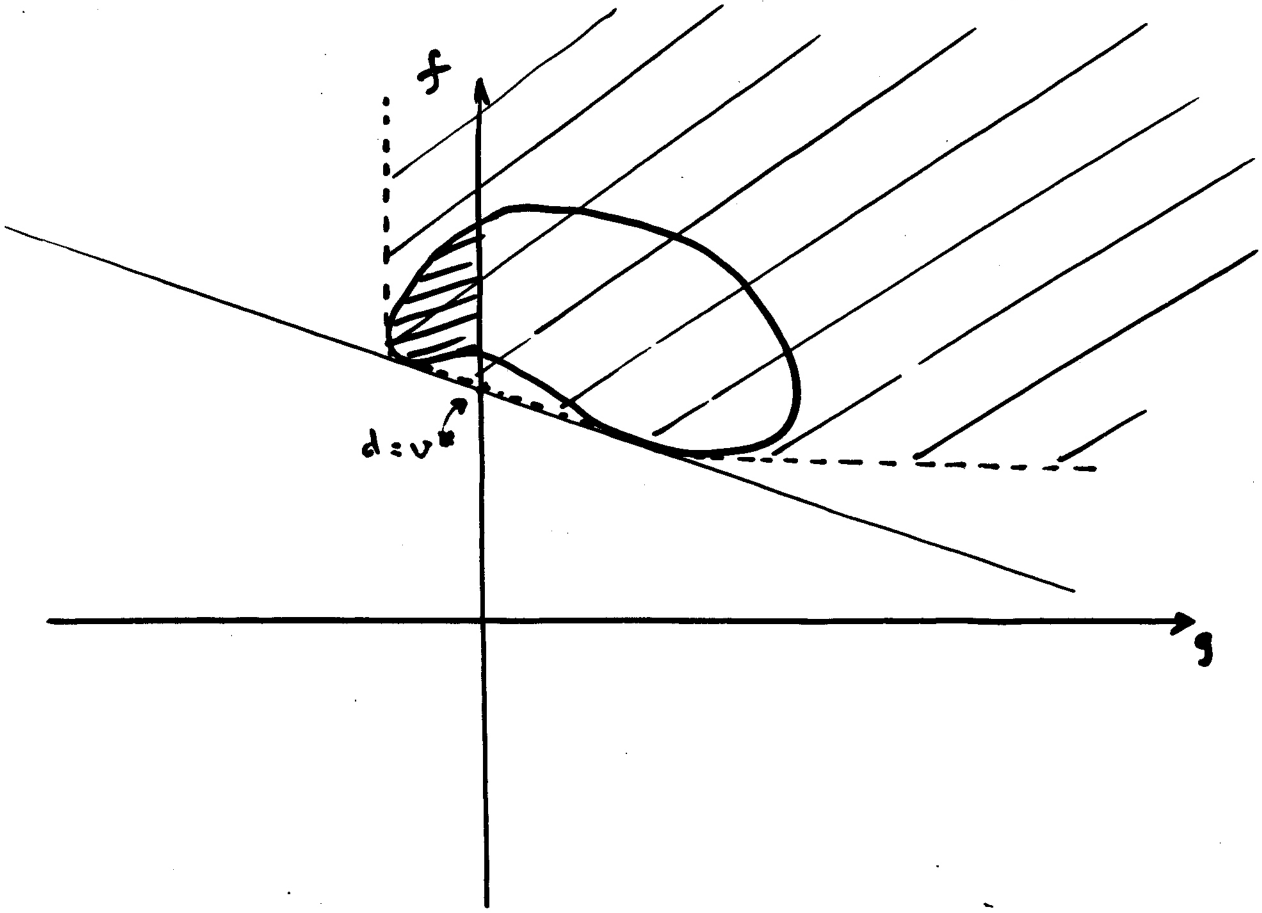
Plano (f,g): Representação de L(u)



Plano (f,g): Problema não-convexo
GAP de dualidade

D6

$$(P^*) \qquad v^* = \min f(x)$$
$$\text{suj. a} \quad g(x) \leq 0$$
$$x \in [X]^c$$



Plano (f, g) : Problema convexificado

DUALIDADE EM PROGRAMAÇÃO INTEIRA

Relaxação Lagrangeana : - Held, Karp 1970, 71
Caixeiro Viajante
- Geoffrion 1974
- Fisher 1981, 85
- Shapiro 1979

$$(P) \quad v = \min_{x \geq 0} \; cx$$

$$Ax \geq b, \; Bx \geq d$$

$$x_j \text{ inteiro}, \; j \in I$$

onde b, c, d são vetores
A, B " matrizes
I conj. que indica as variáveis que devem ser inteiras

Relaxação Lagrangeana : Para $\lambda \geq 0$

$$(PR_\lambda) \quad L(\lambda) = \min_{x \geq 0} \; cx + \lambda(b - Ax)$$

suj. a $Bx \geq d$, x_j inteiro, $j \in I$

Melhor escolha para λ é sol. ótima do dual

$$(D) \qquad d = \max_{\lambda \geq 0} L(\lambda)$$

Obs.: (P) é não-convexo, é provável existir gap de dualidade : $d < v$.

Definindo :

$$(P^*) \qquad v^* = \min \; cx$$
$$\text{suj. a} \quad Ax \geq b$$
$$x \in [X]^c = [\{x \geq 0, Bx \geq d \text{ e } x_j \text{ inteiro}, j \in I\}]^c$$

Qual a relação entre : (P), (PR_λ), (D), (P^*) e ($\bar{P}$) (relaxação usual de PL) ?

Ex.:

$$v = \min \; 7x_1 + 10x_2$$
$$(P) \qquad \text{suj. a} \quad 3x_1 + 5x_2 \geq 7$$
$$0 \leq x_1, x_2 \leq 2$$
$$x_1, x_2 \in \{0,1\}$$

$$(PR_\lambda) \qquad L(\lambda) = \min\left[7\lambda + \{(7-3\lambda)x_1 + (10-5\lambda)x_2\}\right]$$

x_1	x_2	se
0	0	$0 \leq \lambda \leq 2$
0	1	$2 \leq \lambda \leq 7/3$
1	1	$7/3 \leq \lambda$

obtendo-se

$L(\lambda)$	se
7λ	$0 \leq \lambda \leq 2$
$2\lambda + 10$	$2 \leq \lambda \leq 7/3$
$-\lambda + 17$	$7/3 \leq \lambda$

Ex.: $v = \min \; 7x_1 + 10x_2$

(P) s.a. $3x_1 + 5x_2 \geq 7$

$0 \leq x_1, x_2 \leq 2$

$x_1, x_2 \in \{0,1\}$

(PR_λ) $L(\lambda) = \min \{7\lambda + [(7-3\lambda)x_1 + (10-5\lambda)x_2]\}$

$0 \leq x_1, x_2 \leq 2$

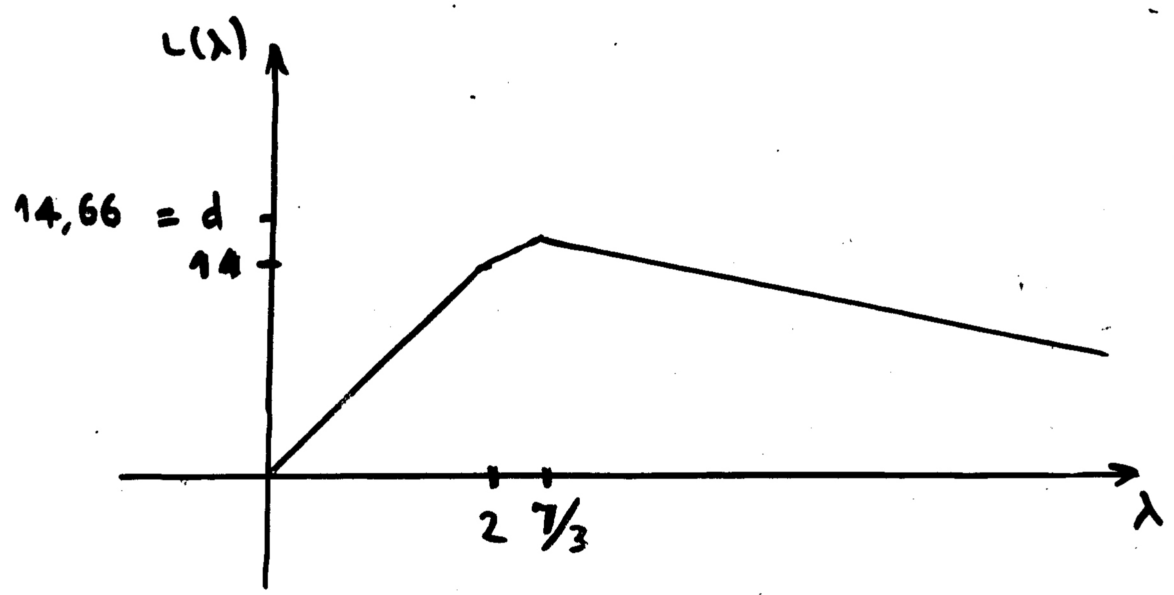
$x_1, x_2 \in \{0,1\}$

Solução:

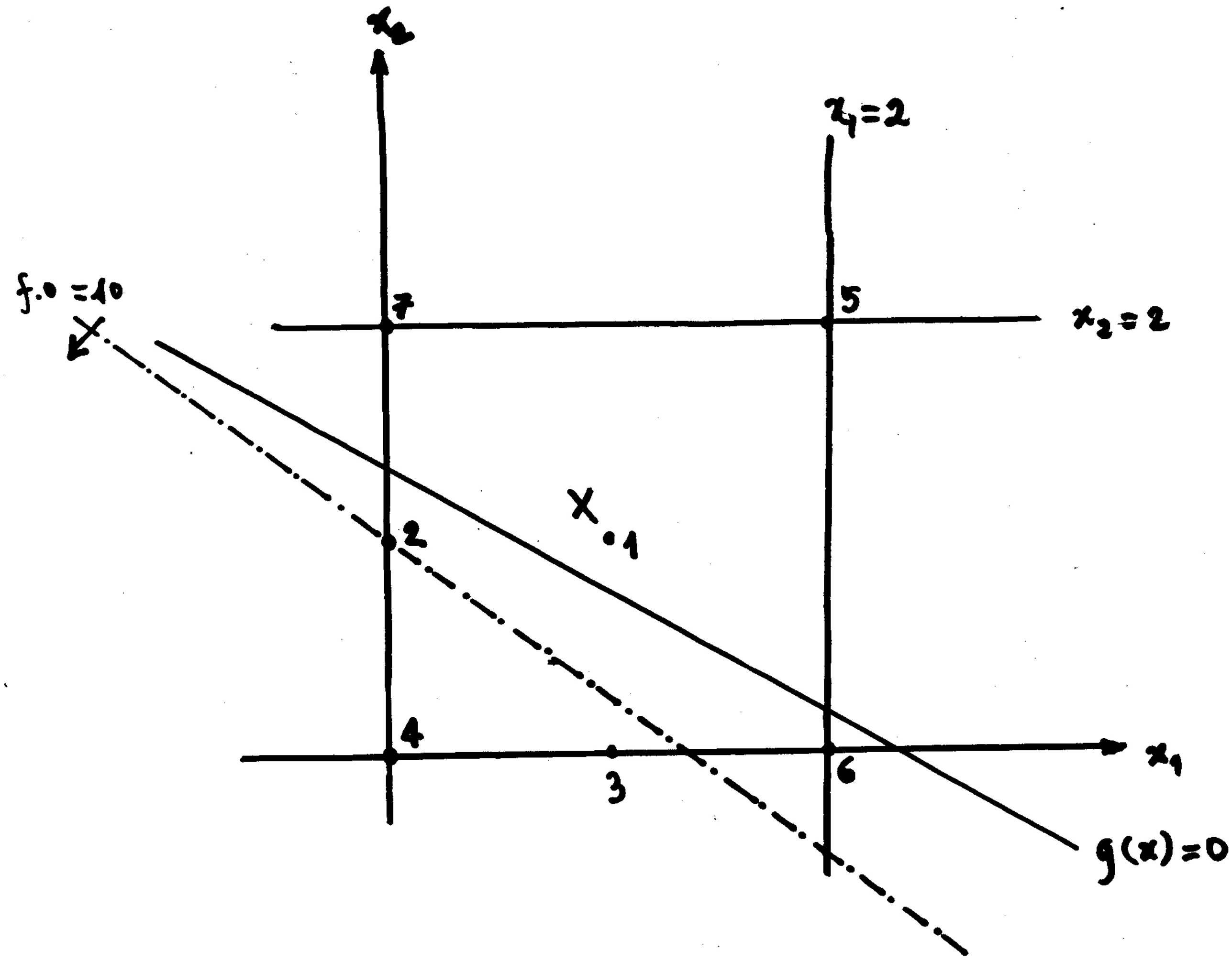
x_1	x_2	se
0	0	$0 \leq \lambda \leq 2$
0	1	$2 \leq \lambda \leq 7/3$
1	1	$7/3 \leq \lambda$

obtendo-se

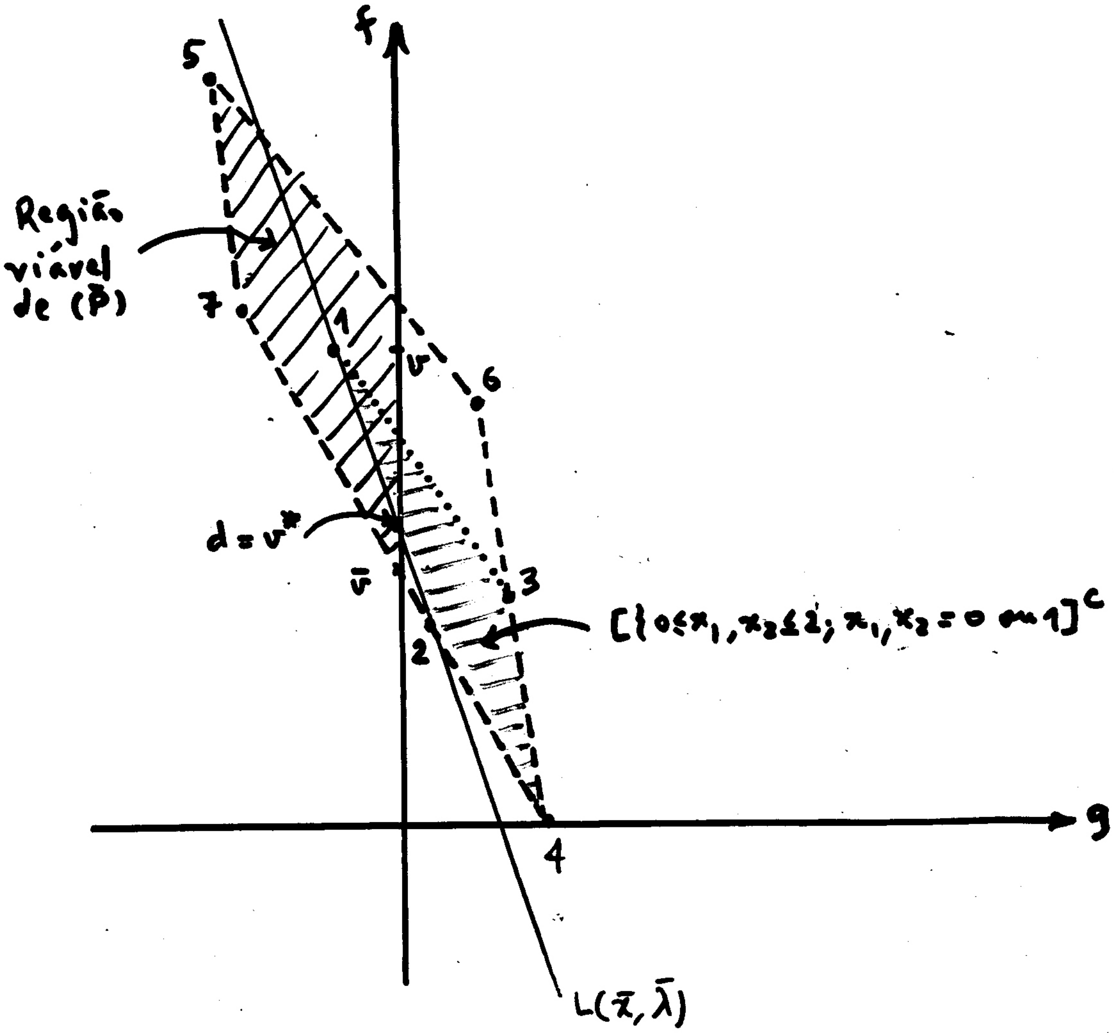
$L(\lambda)$	se
7λ	$0 \leq \lambda \leq 2$
$2\lambda + 10$	$2 \leq \lambda \leq 7/3$
$-\lambda + 17$	$7/3 \leq \lambda$



D10



x_2
$x_1 = 2$
f.o = 10
7
5
$x_2 = 2$
X
1
2
4
3
6
x_1
g(x)=0



f
5
Região viável de ($\bar{P}$)
7
1
v
6
d = v*
$\bar{v}$
3
2
$[\{0 \le x_1, x_2 \le 2;\ x_1, x_2 = 0$ ou $1]^c$
g
4
$L(\bar{x}, \bar{\lambda})$

Uso da Relaxação Lagrangeana

Fisher, 1981

- Como selecionar um valor apropriado para λ?
- Como encontrar um valor de λ tal que d seja igual ou próximo a v?
- Como escolher entre relaxações lagrangeanas diferentes de um mesmo problema?
- Como (PR_λ) pode ser usado para obtenção de soluções viáveis para (P)?
- Como os limites inferiores encontrados pelo problema lagrangeano podem ser usados com o "Branch & Bound"?

EXEMPLO: GAP = Generalized assignment problem (Problema generalizado de atribuição)

$$z = \min \sum_{i=1}^{m} \sum_{j=1}^{n} c_{ij} x_{ij} \qquad (1)$$

suj. a

$$\sum_{i=1}^{m} x_{ij} = 1 \ , \ j = 1, \ldots, n \qquad (2)$$

$$\sum_{j=1}^{n} a_{ij} x_{ij} \leq b_i \ , \ i = 1, \ldots, m \qquad (3)$$

$$x_{ij} = 0 \text{ ou } 1 \ , \ \forall i, j \qquad (4)$$

$$(PR_{\lambda}^{(1)}) \quad z_D(\lambda^{(1)}) = \min \sum_{i=1}^{m} \sum_{j=1}^{n} c_{ij} x_{ij} + \sum_{j=1}^{n} \lambda_j^{(1)} \Big(\sum_{i=1}^{m} x_{ij} - 1 \Big)$$

suj. a (3) e (4)

$$= \min \sum_{i=1}^{m} \sum_{j=1}^{m} (c_{ij} + \lambda_j^{(1)}) x_{ij} - \sum_{j=1}^{n} \lambda_j^{(1)}$$

suj. a (3) e (4)

Este problema se reduz a m problemas da mochila 0-1 e pode ser resolvido em tempo proporcional a $n \sum_{i=1}^{m} b_i$.

$$(PR_{\lambda^{(2)}}) \quad z_D(\lambda^{(2)}) = \min \sum_{i=1}^{m}\sum_{j=1}^{n} c_{ij}x_{ij} + \sum_{i=1}^{m} \lambda_i^{(2)}\left(\sum_{j=1}^{n} a_{ij}x_{ij} - b_i\right)$$

Suj. a (2) e (4)

$$= \min \sum_{j=1}^{n}\left(\sum_{i=1}^{m}(c_{ij} + \lambda^{(2)} a_{ij})x_{ij}\right) - \sum_{i=1}^{m}\lambda_i^{(2)} b_i$$

Suj. a (2) e (4)

Este problema é definido para $\lambda^{(2)} \geq 0$, condição necessária para que $z_D(\lambda^{(2)}) \leq z$. As restrições (2) são limites superiores generalizados (generalized upper bounds = GUB). O problema é de fácil solução em tempo proporcional a mn determinando $\min_i (c_{ij} + \lambda_j^{(2)} a_{ij})$ para cada j e fazendo o valor de $x_{ij} = 1$ para as variáveis associadas e zero para as demais.

$\Rightarrow$ $(PR_{\lambda^{(2)}})$ possui a propriedade de integralidade e $(PR_{\lambda^{(1)}})$ não $\Rightarrow$ $z_D(\lambda^{(1)}) \geq z_D(\lambda^{(2)}) = \overline{Z}$.

DETERMINAÇÃO DE λ

Melhor escolha:

$$(D) \qquad d = \max_{\lambda \geq 0} L(\lambda)$$

Propriedades:

Assumindo $X = \{ x : Bx \geq d, x \geq 0 \text{ e inteiro} \}$ finito

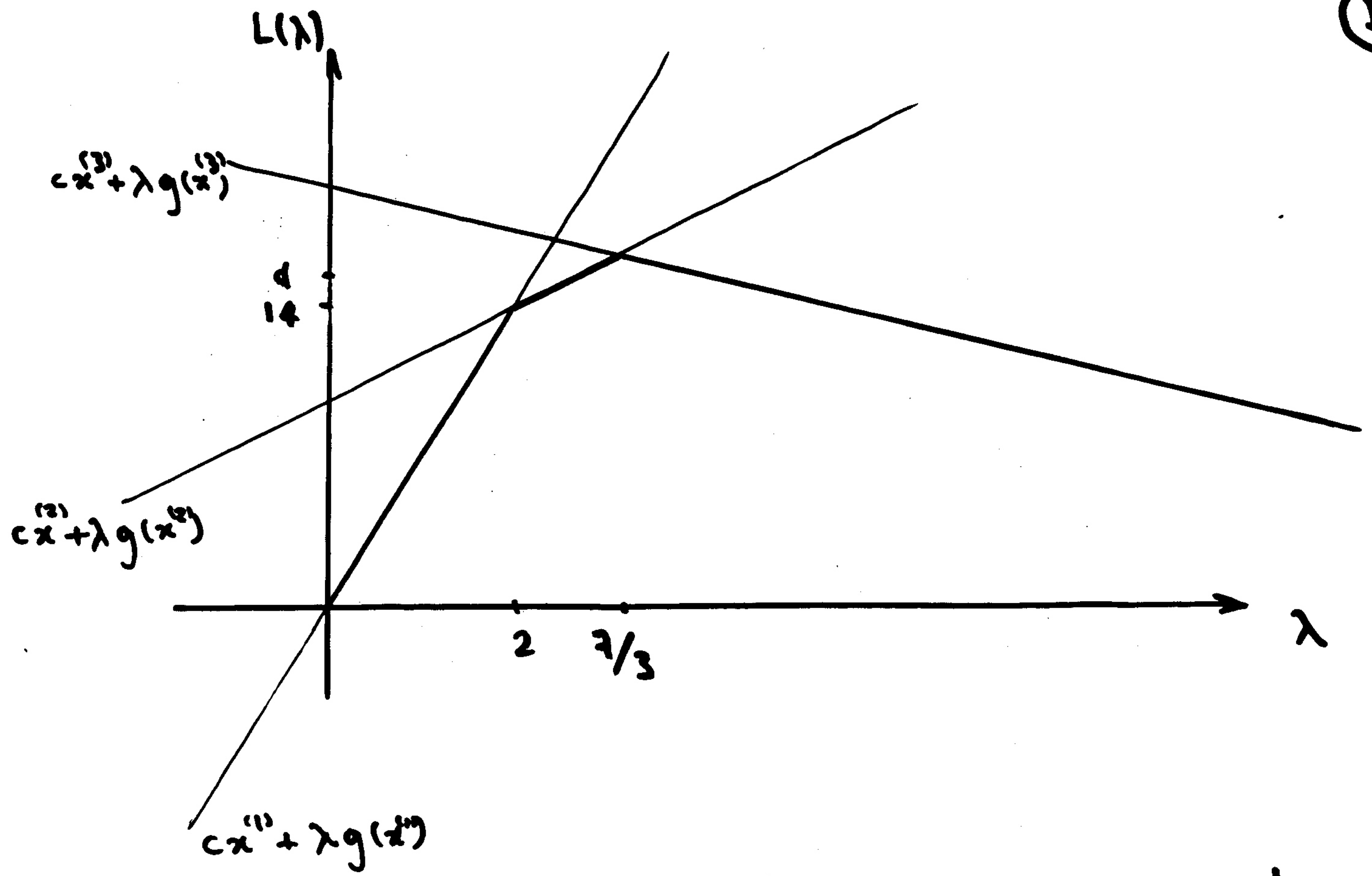
$X = \{ x^t : t = 1, \dots, T \}$, $T = \#$ sol. viáveis para (PR_λ)

Reescrevendo (D)

$$(\tilde{D}) \qquad d = \max_{\lambda \geq 0} w$$

$$\text{suj. a} \quad w \leq cx^t + \lambda(b - Ax^t) = cx^t + \lambda g(x^t), \quad t = 1, \dots, T$$

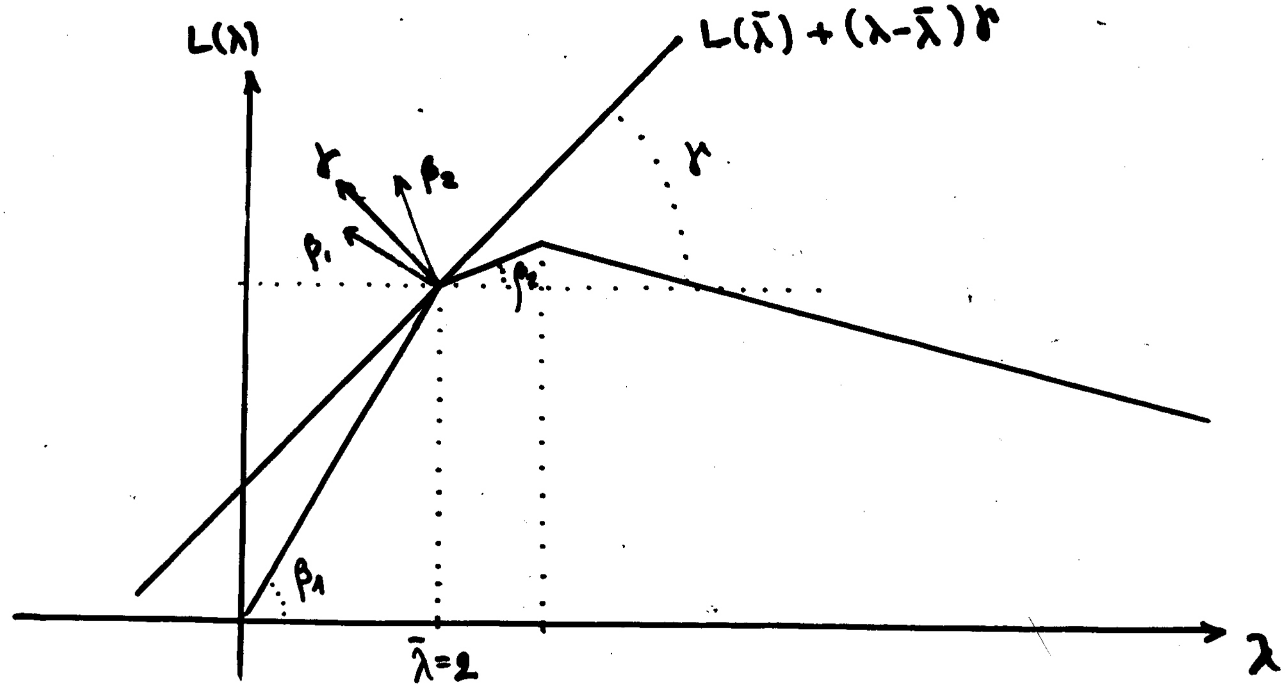
obs.: $L(\lambda)$
- côncava
- contínua
- mas não é diferenciável em todos seus pontos, em geral em um ponto ótimo.



L(λ) para o exemplo

subgradiente (vetor γ) de $L(\lambda)$ em $\bar{\lambda}$:

$$L(\lambda) \leq L(\bar{\lambda}) + (\lambda - \bar{\lambda})\gamma \, , \; \forall \lambda$$



obs.:

- Qualquer subgradiente γ em $\bar{\lambda}$ poderá ser escrito como

$$\gamma = \alpha_1 \beta_1 + \alpha_2 \beta_2 \quad ; \quad \alpha_1, \alpha_2 \geq 0 \quad , \quad \alpha_1 + \alpha_2 = 1 \qquad \left(\begin{array}{c}\text{combinação}\\ \text{convexa}\end{array}\right)$$

- $g(x^t)$ é subgradiente de $L(\lambda)$ em λ^t:

$$L(\lambda) \leq f(x^t) + \lambda g(x^t) = L(\lambda^t) - \lambda^t g(x^t) + \lambda g(x^t) =$$

$$= L(\lambda^t) + (\lambda - \lambda^t)\, g(x^t) \, ,$$

onde $L(\lambda^t) = f(x^t) + \lambda^t g(x^t) = \min\limits_{x \in X} \{ f(x) + \lambda^t g(x) \}$

$\Rightarrow$ <u>Para o exemplo</u>

i	para x_1^i	x_2^i	λ^i	[illegible] $g(x^i)$:
[illegible]	0	0	$0 \leq \lambda \leq 2$	[illegible]
[illegible]	0	1	$2 \leq \lambda \leq 7/3$	[illegible]
3	1	1	$7/3 \leq \lambda$	-1

Subdiferencial: (conj. de todos subgradientes em $\bar{\lambda}$)

$$\partial L(\bar{\lambda}) = \{ \gamma : \gamma = \sum_{i \in T(\bar{\lambda})} \alpha_i \gamma^i \, , \sum_{i \in T(\bar{\lambda})} \alpha_i = 1 \, , \, \alpha_i \geq 0 \, , \forall_i \in T(\bar{\lambda}) \}$$

onde $T(\bar{\lambda}) = \{ i : L(\bar{\lambda}) = L(x^t, \bar{\lambda}) \}$

$\Rightarrow$ - $\bar{\lambda}$ é ótima se e só se $0 \in \partial L(\bar{\lambda})$

- Subdiferencial é um conjunto não-vazio, fechado, limitado e convexo, de $\mathbb{R}^m$.

Método de subgradientes

Dado λ^0, uma sequencia $\{\lambda^k\}$ é gerada usando

$$\lambda^{k+1} = \lambda^k + \alpha_k (b - Ax^k) = \lambda^k + \alpha_k\, g(x^k)$$

onde x^k é uma solução ótima de (PR_{λ^k})

α_k um escalar positivo ("passo")

Fisher, 1981
Held, Wolfe, Crowder, 197
Goffin (1977)

"Tamanho do passo"

$$\alpha_k = \theta_k \frac{(L^* - L(\lambda^k))}{\| g(x^k) \|^2}$$

onde θ_k satisfaz : $0 < \varepsilon_1 \leq \theta_k \leq 2 - \varepsilon_2$, $\varepsilon_2 > 0$;

L^* é um limite superior para $L(\lambda)$ (Heurística)

Poljack (1967) mostrou que $L(\lambda^k) \to d$ se $\alpha_k \to 0$ e $\sum_k \alpha_k \to \infty$. Sugestão: $\theta_0 = 2$ e usar $1/2\, \theta_k$ sempre que $L(\lambda)$ não crescer em um nº pre-determinado de iterações

D18

Adaptação:

$$\lambda_i^{k+1} = \max \{0 , \lambda_i^k + \alpha_k g_i(x^k) \} , \quad i = 1, \dots, m.$$

Para o exemplo:

$L^* = \bar{v} = 10$

i	0	1	2	3	4
λ^i	0	20	7	7	7
θ_i	2	1	0,5		
$L(\lambda^i)$	0	-3	10	10	10

$L^* = 16 > \bar{v}$

i	0	1	2	3	4	5	6	7	8
λ^i	0	32	1	10	1	5,5	3,25	2,125	3
θ_i	2	1	1	1	0,5	0,5	0,5	0,5	0,5
$L(\lambda^i)$	0	-15	7	7	7	11,5	13,75	14,25	14

9	10	11	12	13	14	15	16	17	18	19	20
2	3	2,5	2,125	2,59	2,19	2,59	2,39	2,2	2,4	2,225	2,4
0,5	0,25	0,25	0,25	0,25	0,25	0,25	0,25	0,25	0,25	0,25	0,25
14	14	14,5	14,25	14,41	14,39	14,41	14,61	14,4	14,6	14,45	14,6

$L^* = v^* = d = 14{,}66$

i	0	1	2	3
λ^i	0	29,32	2,33	2,33
θ_i	2	1	1	1
$L(\lambda^i)$	0	12,32	14,66	14,66

Bibliografia:

Fisher, M.L. (1981) "The Lagrangean Relaxation method for Solving Integer Programming Problems" Man. Science

Geoffrion, A.M. (1971) "Duality in Nonlinear Programming" SIAM Review

Geoffrion, A.M. (1974) "Lagrangean Relaxation for Integer Programming" Math. Prog. Study 2

Goffin, J.L. (1977) "On the convergence Rates of Subg. Op. Meth." Mathematical Programming

Held, Karp (1970) "The Traveling Salesman Problem and Minimum Spanning Trees" Op. Research
(1971) Part II

Held, Wolfe, Crowder (1974) "Validation of Subgradient Opt." Math. Prog.

Poljack, B.T. (1967) "A general method for Solving Extremum Problems" Soviet Math. Doklady

Rockafellar, R.T. (1970) "Convex Analysis" Princeton U. Press

Shapiro, J.F. (1979) "Mathematical Programming - Structures and Algorithms" John Wiley

Shapiro, J.F. (1979) "A Survey of Lagrangean techniques for discrete optimization" Annals of Discrete Math.