Condições de otimalidade para problemas com restrições

Departamento de Matemática Universidade do Minho

DMAT-UM 1 / 36

Outline

1 Otimização com restrições de igualdade

2 Otimização com restrições de desigualdade

3 Otimização com restrições de igualdade e desigualdade

4□ > 4□ > 4□ > 4□ > 4□ > 4□

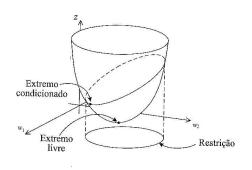
2 / 36

DMAT-UM

Otimização com restrições de igualdade

minimizar
$$F(w)$$

 $w \in \mathbb{R}^d$ $c_n(w) = 0, \quad n \in \mathcal{E} = \{1, \dots, N\}$ (1)



Objetivo: Encontrar um ponto w^* que verifica a restrição $g(w^*) = 0$ e tal que o valor da função F(w) nos outros pontos w (que verificam a restrição g(w) = 0), não é menor do que $F(w^*)$.

Redução a um problema de extremos livres com N=1

Nota: Assume-se que F e c_1 são continuamente diferenciáveis até à $2^{\underline{a}}$ ordem.

A situação mais simples é quando a equação

$$c_1(w) = 0 \Leftrightarrow c_1(w_1, w_2, \cdots, w_N) = 0,$$

que define a restrição, permite explicitar facilmente uma das varáveis como função das outras, por exemplo,

$$w_N = \Phi(w_1, \cdots, w_{N-1}).$$

Neste caso, substituindo w_n pele sua expressão como função de w_1, \dots, w_{n-1} na expressão F,

$$F(w_1, \dots, w_{N-1}, \Phi(w_1, \dots, w_{N-1})).$$

Este problema reduz-se a um problema de optimização sem restrições.

Exercício: Uma caixa paralelipipéda sem tampa tem 32 cm³ de volume. Determine as suas dimensões de maneira a minimizar o custo do seu fabrico (que se supõe proporcional à área das suas faces).

Exercício: Determina a distância mais curta do ponto (2,1,-1) ao plano $w_1 + w_2 - w_3 = 1$.

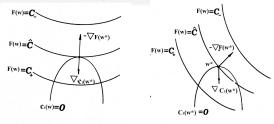
Nem sempre na restrição se pode explicitar uma das variáveis como função da outra.

Otimização com restrições de igualdade com N=1

Recordar:

- Se $\nabla F(w^*) \neq 0$, então $\nabla F(w^*)$ é normal à curva de nível de F em w^* .
- Se $\nabla c_1(w^*) \neq = 0$, então $\nabla c_1(w^*)$ é normal à curva de nível de c_1 em w^* .
- Nestes casos, as duas curvas de nível são tangentes em w* se e só se $\nabla F(w^*)$ e $\nabla c_1(w^*)$ têm a mesma direção.
- O minimizante é encontrado quando as curvas de nível são tangentes, caso contrário é possível encontrar pontos no conjunto $\{w:c_1(w)=0\}$, onde o valor da função F é estritamente menor do que no ponto w^* .

Otimização com restrições de igualdade com N=1



Esta observação geométrica pode ser transformada num teorema, chamado método dos Multiplicadores de Lagrange:

Teorema 1

Existem escalares λ_0 e λ_1 (multiplicadores de Lagrange) tais que:

•
$$\lambda_0 \nabla F(w^*) + \lambda_1 \nabla c_1(w^*) = 0$$

•
$$\lambda_0^2 + \lambda_1^2 \neq 0$$

Nota: A segunda condição é para garantir que a primeira não seja satisfeita com λ_0 e λ_1 iguais a zero.

Pretende-se fortalecer as condições na forma

$$\nabla F(w^*) + \lambda_1 \nabla c_1(w^*) = 0, \text{ isto } \epsilon \lambda_0 = 1$$
 (2)

garantido que estas selecionam os candidatos a ótimo.

Esta afirmação só é verdadeira se $\nabla c_1(w^*) \neq 0$.

Por exemplo: Se $\nabla c_1(w^*)=0$, $c_1(w^*)=0$ e $c_1(w)>0$, $w\neq w^*$, então $\mathcal{D}=\{w^*\}$.

Para qualquer função F o ponto w^* é um minimizante, em particular uma função com $\nabla F(w^*) \neq 0$, neste caso a equação (2) não é satisfeita.

Otimização com várias restrições de igualdade

Considere-se várias restrições do tipo de igualdade $c_n(w)=0$, com $n\in\mathcal{E}=\{1,\ldots,N\}$, então o vetor é normal à superfície $\mathcal{D}=\{w:c_n(w)=0,n\in\mathcal{E}=\{1,\ldots,N\}\}$, no minimizante w^* é uma combinação linear dos gradientes

$$\sum_{n=1}^{N} \lambda_n \nabla c_n(w^*).$$

O método dos multiplicadores de Lagrange neste caso tomam a seguinte forma:

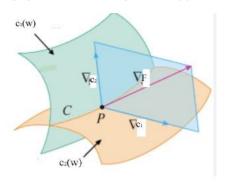
Teorema 2

Existem escalares λ_0 e λ_n com $n \in \mathcal{E} = \{1, \dots, N\}$ tais que:

•
$$\lambda_0 \nabla F(w^*) + \sum_{n=1}^N \lambda_n \nabla c_n(w^*) = 0$$

•
$$\lambda_0^2 + \sum_{n=1}^N \lambda_n^2 \neq 0$$

Geometricamente, basta substituir a superfície $\{w: c_1(w)=0\}$, pela superfíce $\mathcal{D}=\{w: c_n(w)=0, n\in\mathcal{E}=\{1,\ldots,N\}\}$,



Novamente, se $\lambda_0=0$, no Teorema 2, as condições são uma relação entre restrições em que a função a minimizar não entra nas condições.

São necessárias hipóteses adicionais (qualificação de restrição ou ponto regular) para garantir que o problema satisfaz as condições de otimalidade com $\lambda_0=1$.

.

Definição 3 (Ponto Regular para o conjunto definido por restrições de igualdade)

Seja $\mathcal{D}=\{w:c_n(w)=0,n\in\mathcal{E}=\{1,\ldots,N\}\}$. Um ponto w^* é um ponto regular se os vetores $\nabla c_n(w^*)$, $n\in\mathcal{E}=\{1,\ldots,N\}$ são linearmente independentes.

Exercício: Considere as restrições definidas por: $w_1^2 + w_2^2 + w_3^2 - 3 = 0$ e $2w_1 - 4w_2 + w_3^2 + 1 = 0$. Verifique se o ponto $w^* = (1, 1, 1)^T$ é admissível e regular.

As condições de otimalidade também podem ser estabelecidas em termos da Função Lagrangiana.

Definição 4

A função Lagrangiana é definida pela função objetivo F e pelas funções de restrição c_n . A Função Lagrangiana associada ao problema (1) é

$$L(w,\lambda) = F(w) - \sum_{n=1}^{N} \lambda_n c_n(w)$$

onde λ_n é designado por multiplicador de Lagrange associada à restrição $c_n(w) = 0 \ (n = 1, ..., N)$.

Notação:

 $\lambda = (\lambda_1, \dots, \lambda_N)^T$ é vetor dos multiplicadores de Lagrange.



Condições de Karush-Kuhn-Tucker (KKT)

Teorema 5 (Condição necessária de 1ª ordem de KKT)

Seja w^* um minimizante local do problema (1). Se w^* é um ponto regular das restrições então existe um vetor de multiplicadores de Lagrange λ^* tal que as seguintes condições são satisfeitas em (w^*, λ^*) :

$$abla_w L(w^*,\lambda^*)=0$$
 condição de 1ª ordem $c_n(w^*)=0$ admissibilidade $(n=1,\ldots,N)$

- $\nabla_w L(w^*, \lambda^*) = 0 \Leftrightarrow \nabla F(w^*) = \sum_{n=1}^N \lambda_n^* \nabla c_n(w^*)$
- \bullet (w^*, λ^*) é ponto estacionário da Lagrangiana e chama-se ponto KKT.
- encontrar pontos estacionários do problema (1)
 ⇔ encontrar pontos estacionários da Lagrangiana

Exercício: Determine a solução para os seguintes problemas:

b) $w \in \mathbb{R}^2$ 1 2 sujeito a $w_1^2 + w_2^2 - 1 = 0$,

Exercício: Considere o seguinte problema:

$$\begin{array}{ll} \underset{w \in \mathbb{R}^2}{\text{minimizar}} & -w_1^2 \\ \text{sujeito a} & (1-w_1)^3 + w_2 = 0, \\ & (1-w_1)^3 - w_2 = 0. \end{array}$$

Mostre que $w^* = (1,0)$ é o único minimizante global, contudo este ponto não satisfaz as condições necessárias de KKT. Justifique.

Teorema 6 (Condições necessárias de 2ª ordem de KKT)

Seja w^* um minimizante local do problema (1) que é ponto regular. Seja λ^* o vetor de multiplicadores de Lagrange que verifica as condições KKT. Então:

$$s^T \nabla^2_{ww} L(w^*, \lambda^*) s \ge 0, \ \forall s \in \mathcal{C},$$

onde
$$C = \{s : \nabla c_n(w^*)^T s = 0, \forall n \in \{1, ..., N\}\}.$$

Observação:

A matriz hessiana da função Lagrangiana em ordem a w é

$$\nabla^2_{w \, w} L(w^*, \lambda^*) = \nabla^2 F(w^*) - \sum_{n=1}^N \lambda_n^* \nabla^2 c_n(w^*)$$

• O conjunto C é constituído por direções tangentes à curvas de niveis $c_n(w^*) = 0, \forall n \in \{1, ..., N\}.$

Teorema 7 (Condições suficientes de 2^a ordem de KKT)

Suponha que para algum ponto admissível w^* do problema (1) existe um vetor de multiplicadores de Lagrange λ^* que verifica as condições KKT. Suponha também que para

$$s^T \nabla^2_{ww} L(w^*, \lambda^*) s > 0, \ \forall s \in \mathcal{C},$$

onde
$$C = \{s \neq 0 : \nabla c_n(w^*)^T s = 0, \forall n \in \{1, \dots, N\}\}.$$

Então w* é minimizante local estrito do problema (1).

Exercício: Verifique que as condições suficientes de $2^{\underline{o}}$ ordem são satisfeitas para $\begin{array}{ccc} & \min & w_1 + w_2 \\ & w \in \mathbb{R}^2 \\ & \text{sujeito a} & w_1^2 + w_2^2 - 2 = 0, \end{array}$

Exercício: Determine a solução para o seguintes problema:

Otimização com restrições de desigualdade

minimizar
$$F(w)$$

 $w \in \mathbb{R}^d$
sujeito a $c_n(w) \ge 0$, $n \in \mathcal{I} = \{1, \dots, N\}$ (3)

Observação:

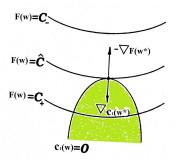
Se no ponto w^* as funções

$$c_n(w*) > 0, \forall n \in \mathcal{I} = \{1, \ldots, N\}$$

então todos os pontos de uma vizinhança de w^* verificam a restrição, o problema (pelo menos localmente) torna-se um problema sem restrições e as condições de otimalidade para problemas sem restrições podem ser aplicadas.

Otimização com restrições de desigualdade com N=1

Seja w^* um ponto que verifica a condição $c_1(w^*)=0$ e tal que o valor da função F(w), nos outros pontos w que verifica, $c_1(w)\geq 0$, não é menor que $F(w^*)$. Isto só é possível se os gradientes das funções F e c_1 no ponto w^* tiverem a mesma direção de acordo com a ilustração abaixo.



Caso contrário, em qualquer vizinhança do ponto w^* , existem pontos tais que $c_1(w) > 0$ e $F(w) < F(w^*)$.

Analiticamente, temos as condições necessárias de otimalidade escritas na seguinte forma:

Teorema 8

Existem escalares $\lambda_0 \geq 0$ e $\lambda_1 \geq 0$ (multiplicadores de Lagrange) tais que:

- $\lambda_0 \nabla F(w^*) + \lambda_1 \nabla c_1(w^*) = 0$
- $\lambda_1 c_1(w^*) = 0$ (condição complementar)
- $\lambda_0 + \lambda_1 \neq 0$

Observação: As duas últimas condições mostram que se $c_1(w^*) > 0$, o multiplicador $\lambda_1 = 0$ e método do multiplicadores de Lagrange transformam-se na $1^{\underline{o}}$ condição de otimalidade.

Otimização com N restrições de desigualdade

Generalizando o teorema (8) para um problema com N restrições de desigualdade obtém-se:

Teorema 9

Existem escalares $\lambda_n \geq 0$ com $n \in \{0, \dots, N\}$ (multiplicadores de Lagrange) tais que:

•
$$\lambda_0 \nabla F(w^*) + \sum_{n=1}^N \lambda_n \nabla c_n(w^*) = 0$$

•
$$\lambda_n c_n(w^*) = 0, \forall n \in \{1, ..., N\}$$

$$\bullet \ \sum_{n=0}^{N} \lambda_n \neq 0$$

Observação: Novamente, se $\lambda_0=0$ estas condições são uma mera relação entre as restrições e a função que se pretende minimizar não aparece nas condições.

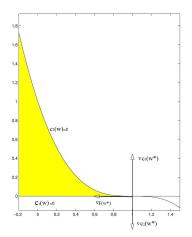
O exemplo abaixo mostra que nem sempre é possível escolher $\lambda_0=1$. **Exemplo:**

$$\begin{array}{ll} \underset{w \in \mathbb{R}^2}{\text{minimizar}} & -w_1 \\ \text{sujeito a} & -w_2 - (w_1 - w_2)^3 \geq 0, \\ & w_1 \geq 0, \end{array}$$

A solução é $w^* = (w_1, w_2) = (1, 0)$. Contudo a condição

$$\lambda_0 \nabla F(w^*) + \sum_{n=1}^2 \lambda_n \nabla c_n(w^*) = 0$$

é apenas satisfeita para $\lambda_0=0$



Observação: Novamente, é necessário identificar os problemas em condições necessárias de otimalidade possam ser escritas com $\lambda_0 = 1$.

Nota: Neste problema os vetores $\nabla c_1(w^*)$ e $\nabla c_2(w^*)$ são linearmente dependentes.

Definição 10 (restrição ativa ou não ativa)

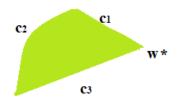
Seja $w^* \in \mathcal{D}$ um ponto admissível. Uma restrição de desigualdade, $c_n(w) \geq 0$, é dita ativa em w^* , se $c_n(w^*) = 0$. Caso $c_n(w^*) > 0$, diz-se que c_n é não ativa em w^* .

Definição 11 (conjunto ativo num ponto admissível)

O conjunto ativo num ponto admissível é conjunto dos índices das restrições de desigualdade ativas no ponto admissível w^* :

$$\mathcal{A}(w^*) = \{n : c_n(w^*) = 0, n \in \mathcal{I}\}.$$

Exemplo:



$$\mathcal{A}(w^*) = \{1, 3\}.$$

Definição 12 (Ponto Regular para o conjunto definido por restrições de desigualdade)

Seja $w^* \in \mathcal{D}$ um ponto admissível e $\mathcal{A}(w^*)$ o conjunto ativo em w^* . O ponto admissível é designado por ponto regular se o conjunto dos gradientes das restrições ativas em w^* , $\{\nabla c_n(w^*) : n \in \mathcal{A}(w^*)\}$, é linearmente independente.

Exercício: Considere as restrições definidas por: $w_1 \ge 0$, $w_2 \ge 0$ e $-(w_1-1)^2+w_2 \ge 0$. Verifique se o ponto $w^*=(1,0)^T$ é admissível e regular.

As condições de otimalidade novamente podem ser estabelecidas em termos da Função Lagrangiana.

A Função Lagrangiana associada ao problema (3), é definida por

$$L(w,\lambda) = F(w) - \sum_{n=1}^{N} \lambda_n c_n(w)$$

onde λ_n é o multiplicador de Lagrange associada à restrição $c_n(w) \geq 0$ (n = 1, ..., N).

Notação:

 $\lambda = (\lambda_1, \dots, \lambda_N)^T$, é vetor dos multiplicadores de Lagrange.

Condições de Karush-Kuhn-Tucker (KKT)

Teorema 13 (Condição necessária de 1ª ordem de KKT)

Seja w^* um minimizante local do problema (3). Se w^* é um ponto regular das restrições então existe um vetor de multiplicadores de Lagrange λ^* tal que as seguintes condições são satisfeitas em (w^*, λ^*)

$$abla_w L(w^*, \lambda^*) = 0$$
 condição $1^{\underline{a}}$ ordem $c_n(w^*) \geq 0$ admissibilidade $(n = 1, \dots, N)$ $\lambda_n^* \geq 0$ admissibilidade $(n = 1, \dots, N)$ $\lambda_n^* c_n(w^*) = 0$ complementaridade $(n = 1, \dots, N)$

Observação:

•
$$\nabla_w L(w^*, \lambda^*) = 0 \Leftrightarrow \nabla F(w^*) = \sum_{n=1}^N \lambda_n^* \nabla c_n(w^*)$$

A condição

$$\lambda_n^* c_n(w^*) = 0 \ (n = 1, \dots, N)$$

- , chama-se condição de complementaridade e significa que:
 - ou a restrição n está ativa na solução $(c_n(w^*) = 0)$
 - ou o multiplicador λ_n^* que lhe está associado é **nulo** $(\lambda_n^* = 0)$
 - Se os multiplicadores que correspondem a restrições ativas são todos positivos, a complementaridade diz-se estrita.
 - Qualquer restrição que **não esteja ativa** em w^* ($c_n(w^*) > 0$) tem multiplicador **nulo**.
 - Se o multiplicador que corresponde a uma restrição ativa é positivo, a restrição diz-se não degenerada.
 - Se um multiplicador que corresponde a uma restrição ativa é nulo, a restrição diz-se degenerada.



Exercício: Considere o seguinte problema:

$$\begin{array}{ll} \underset{w \in \mathbb{R}^2}{\text{minimizar}} & F(w) = w_1 \\ \text{sujeito a} & (w_1 + 1)^2 + w_2^2 \geq 1 \\ & w_1^2 + w_2^2 \leq 2 \end{array}$$

Verifique se os pontos $w^{*1} = (0,0)^T$, $w^{*2} = (-1,-1)^T$, $w^{*3} = (0,\sqrt{2})^T$ satisfazem as condições de otimalidade de $1^{\underline{a}}$ ordem.

Exercício: Determine as soluções para os seguintes problemas:

Teorema 14 (Condições necessárias de 2º ordem de KKT)

Seja w^* um minimizante local do problema (3) que é ponto regular. Seja λ^* o vetor de multiplicadores de Lagrange que verifica as condições KKT e a condição de **complementaridade estrita**. Então

$$s^T \nabla^2_{ww} L(w^*, \lambda^*) s \ge 0, \forall s \in \mathcal{C}(w^*, \lambda^*)$$

onde
$$C(w^*, \lambda^*) = \{ s \in \mathbb{R}^d : \nabla c_n(w^*)^T s = 0, \ \forall n \in \mathcal{A}(w^*) \ com \ \lambda_n^* > 0 \}$$

Nota: O conjunto C é constituído pelos vetores tangentes às superfícies das **restrições ativas e não degeneradas**.

Teorema 15 (Condições suficiente de 2ª ordem de KKT)

Suponha que para algum ponto admissível w^* do problema (3) existe um vetor de multiplicadores de Lagrange λ^* que verifica as condições KKT e a condição de complementariedade estrita. Suponha também que

$$s^T \nabla^2_{ww} L(w^*, \lambda^*) s > 0, \forall s \in \mathcal{C}(w^*, \lambda^*), s \neq 0$$

onde
$$C(w^*, \lambda^*) = \{ s \in \mathbb{R}^d : \nabla c_n(w^*)^T s = 0, \ \forall n \in \mathcal{A}(w^*) \ com \ \lambda_n^* > 0 \}$$

Então w* é minimizante local estrito do problema (3).

Observação:

• $s^T \nabla^2_{w \, w} L(w^*, \lambda^*) s < 0 \Rightarrow w^*$ é maximizante local estrito se se pretende maximizar em vez de minimizar no problema (3).

Otimização com restrições de igualdade e desigualdade

minimizar
$$F(w)$$

sujeito a $c_n(w) = 0$, $n \in \mathcal{E} = \{1, \dots, j\}$
 $c_n(w) \geq 0$, $n \in \mathcal{I} = \{j + 1, \dots, N\}$ (4)

Definição 16 (conjunto ativo num ponto admissível)

O conjunto ativo num ponto admissível é conjunto dos índices das restrições de igualdade e dos índices das restrições de desigualdade ativas no ponto admissível w^* :

$$\mathcal{A}(w^*) = \mathcal{E} \cup \{n \in \mathcal{I} : c_n(w^*) = 0\}.$$

◆ロト 4億ト 4 恵 ト (意) 夕久(*)

DMAT-UM desigualdade 31 / 36

Definição 17 (Ponto Regular para o conjunto definido por restrições de desigualdade)

Seja $w^* \in \mathcal{D}$ um ponto admissível e $\mathcal{A}(w^*)$ o conjunto ativo em w^* . O ponto admissível é designado por ponto regular se o conjunto dos gradientes das restrições ativas em w^* , $\{\nabla c_n(w^*): n \in \mathcal{A}(w^*)\}$, é linearmente independente.

A Função Lagrangiana associada ao problema (4), é definida por

$$L(w,\lambda) = F(w) - \sum_{n \in \mathcal{E} \cup \mathcal{I}} \lambda_n c_n(w)$$

DMAT-UM desigualdade 32 / 36

Condições de Karush-Kuhn-Tucker (KKT)

Teorema 18 (Condição necessária de 1ª ordem de KKT)

Seja w^* um minimizante local do problema (3). Se w^* é um ponto regular das restrições então existe um vetor de multiplicadores de Lagrange λ^* tal que as seguintes condições são satisfeitas em (w^*, λ^*)

$$abla_w L(w^*,\lambda^*)=0$$
 $c_n(w^*)=0, \quad \text{para todo } n\in\mathcal{E}$
 $c_n(w^*)\geq 0, \quad \text{para todo } n\in\mathcal{I}$
 $\lambda_n^*\geq 0, \quad \text{para todo } n\in\mathcal{I}$
 $\lambda_n^*c_n(w^*)=0, \quad \text{para todo } n\in\mathcal{E}\cup\mathcal{I}$

4 D > 4 D > 4 D > 4 D > 9 Q Q

33 / 36

DMAT-UM designaldade

Definição 19

O conjunto critico é definido da seguinte forma:

$$w \in \mathcal{C}(w^*, \lambda^*)$$
 se e só se
$$\left\{ \begin{array}{ll} c_n(w^*)^T w = 0 & \text{para todo } n \in \mathcal{E} \\ \\ c_n(w^*)^T w = 0 & \text{para todo } n \in \mathcal{A}(w^*) \cap \mathcal{I}, \text{ com } \lambda_n^* > 0 \\ \\ c_n(w^*)^T w \geq 0 & \text{para todo } n \in \mathcal{A}(w^*) \cap \mathcal{I}, \text{ com } \lambda_n^* = 0 \end{array} \right.$$

Teorema 20 (Condições necessárias de 2º ordem de KKT)

Seja w^* um minimizante local do problema 4 que é ponto regular. Seja λ^* o vetor de multiplicadores de Lagrange que verifica as condições KKT. Então conjunto dos índices das restrições de desigualdade ativas no ponto admissível w^* :

$$s^T \nabla^2_{w w} L(w^*, \lambda^*) s \geq 0, \forall s \in \mathcal{C}(w^*, \lambda^*).$$

DMAT-UM designaldade 34 / 36

Teorema 21 (Condições suficiente de 2ª ordem de KKT)

Suponha que para algum ponto admissível w^* do problema (3) existe um vetor de multiplicadores de Lagrange λ^* que verifica as condições KKT. Suponha também que

$$s^T \nabla^2_{ww} L(w^*, \lambda^*) s > 0, \forall s \in \mathcal{C}(w^*, \lambda^*), s \neq 0$$

Então w* é minimizante local estrito do problema 4.

Exercício: Determine a solução para o seguintes problema:

minimizar
$$w_1^2 + w_2^2 + w_3^2$$

sujeito a $2w_1 - w_2 + w_3 \le 5$
 $w_1 + w_2 + w_3 = 3$

→ ロト→便ト→度ト→度トー度 のQで

DMAT-UM designaldade 35 / 36

Os apontamentos foram baseados na seguinte bibliografia:[1], [2] e [3].

A. S. Bazaraa, H. Sherali, and C. Shetty.

Nonlinear Programming.

John Wiley and Sons, Inc., 1993.

J. Nocedal and S. J. Wright.

Numerical optimization.

Springer, 1999.

G. Smirnov and V. Bushenkov.

Curso de optimização: programação matemática: cálculo de variações: controlo óptimo.

2005.

DMAT-UM designaldade 36 / 36