Máximos e mínimos para físicos

Henrique Fleming

18 outubro 2001

1 Introdução

Estas notas destinam-se a introduzir, para estudantes de física, o método dos multiplicadores de Lagrange para resolver problemas de máximos e mínimos "relativos", isto é, com condições suplementares. Por exemplo, o problema clássico de construir uma caixa de papelão com a forma de paralelepípedo que tenha máximo volume para uma área de papelão dada, pode ser formulado assim: determinar o máximo da função xyz (volume do paralelepípedo de arestas x, y, z) com a condição de que $xy + xz + yz = K$ (a área total das faces é fixa, e igual a $2K$). Esta condição é denominada *vínculo*, ou *condição subsidiária*.

A solução que ocorre imediatamente é: usar a condição de área fixa para eliminar, digamos, z, em função de x e y. O volume passa extão a ser uma função só de x e y, à qual se aplicam os métodos usuais. O problema se complica, e este método se torna inaplicável, quando as *condições subsidiárias* são muitas ou muito complicadas. Lagrange [1] inventou uma maneira genial de simplificar o problema: o método dos multiplicadores de Lagrange.

Quais são os "métodos usuais" que mencionamos acima? Consistem em procurar os "pontos críticos", ou seja, os pontos em que todas as derivadas parciais da função a maximizar se anulam. Um teorema clássico diz que os pontos em que a função pode ter máximos ou mínimos estão entre esses pontos críticos. Para determinar se, efetivamente, um ponto crítico é ponto de máximo, de mínimo ou nem um nem outro, é preciso examinar as derivadas parciais de ordem mais alta. Não tocaremos neste assunto aqui, limitando-nos a dar boas referências. Os livros de Courant [1], Apostol [2] e Elon Lages Lima [3] são referências de boa qualidade e leitura interessante. Se o leitor tiver acesso, recomendamos especialmente a grande obra clássica de Camille Jordan [4], na qual baseamos o nosso tratamento [2].

[1] Imenso matemático italiano (de Turim) que, com Euler, construiu a análise matemática e a mecânica como as conhecemos hoje, a partir das idéias de Isaac Newton. O primeiro grande trabalho de Lagrange foi explicar por que a Lua mostra sempre a mesma face para a Terra, o que ele fez com cerca de 14 anos de idade! Quanto à mecânica, sua contribuição mais importante foi o chamado *formalismo lagrangeano*, que conduziu ao princípio de mínima ação, em geral associado ao nome de Hamilton, mas já presente na obra de Lagrange. Uma das ruas principais da cidade de Turim se chama *via Lagrange*.

[2] Nenhum elogio é excessivo para essa obra. Três grandes matemáticos atribuem seu gosto

2 O método

2.1 Um exemplo simples

Seja $f(x, y, z)$ a função da qual queremos saber os máximos e mínimos, e seja

$$g(x, y, z) = 0 \tag{1}$$

a condição subsidiária. No problema citado acima, teríamos

$$f(x, y, z) = xyz \tag{2}$$

e $g(x, y, z) = 0$ seria dada por

$$xy + yz + zx - K = 0 \tag{3}$$

Desta última relação segue que

$$z = \frac{K - xy}{x + y} \tag{4}$$

que, levada à Eq.(2), dá:

$$F(x, y) = f(x, y, z(x, y)) = xy\frac{K - xy}{x + y} \tag{5}$$

A solução é obtida agora igualando a zero as derivadas parcias $\frac{\partial F}{\partial x}$ e $\frac{\partial F}{\partial y}$, que são

$$\begin{aligned} \frac{\partial F}{\partial x} &= \frac{y^2}{(x+y)^2}\{K - 2xy - x^2\} \\ \frac{\partial F}{\partial y} &= \frac{x^2}{(x+y)^2}\{K - 2xy - y^2\} \end{aligned}$$

Igualadas a zero, obtemos as equações

$$\begin{aligned} 2xy + x^2 &= K \\ 2xy + y^2 &= K \end{aligned}$$

de onde se conclui que $x = y$ e que $K = 3x^2$. Logo, temos também $x = z$, ou seja, o paralelepípedo de volume máximo, para área dada, é o cubo (existem soluções para o volume mínimo. Quais são?).

pela matemática à leitura de Jordan: Godfrey Hardy, André Weil e Freeman Dyson (mais conhecido como físico teórico, mas também grande matemático, com resultados importantes na Teoria dos Números).

2.2 Multiplicadores de Lagrange

A receita para achar os pontos de máximo é igualar a zero todas as derivadas parciais. Se não houvesse vínculos, isto seria o mesmo que impor $df = 0$, onde df, o diferencial da função f, é dado por

$$df = \frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy + \frac{\partial f}{\partial z}dz \tag{6}$$

Uma vez eliminado z por meio do vínculo, temos, em lugar desta última, a equação

$$dF = \frac{\partial F}{\partial x}dx + \frac{\partial F}{\partial y}dy = 0 \tag{7}$$

ou seja, não aparece mais o diferencial dz, indicando que a função F não depende de z. O método de Lagrange oferece uma técnica mais eficiente e simétrica para eliminar a dependência em z, ou seja, para se livrar do termo em dz na expressão do diferencial da função cujos máximos se procura.

Considere o diferencial da função f:

$$df = \frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy + \frac{\partial f}{\partial z}dz \tag{8}$$

e, como $g(x,y,z) = 0$, temos

$$dg = \frac{\partial g}{\partial x}dx + \frac{\partial g}{\partial y}dy + \frac{\partial g}{\partial z}dz = 0 \tag{9}$$

Sejs λ um número qualquer, de valor a ser determinado posteriormente. Adicionemos a df a quantidade λdg, que é zero. Logo,

$$df = df + \lambda dg$$

Portanto, podemos escrever

$$df = \left(\frac{\partial f}{\partial x} + \lambda\frac{\partial g}{\partial x}\right)dx + \left(\frac{\partial f}{\partial y} + \lambda\frac{\partial g}{\partial y}\right)dy + \left(\frac{\partial f}{\partial z} + \lambda\frac{\partial g}{\partial z}\right)dz \tag{10}$$

Mas, como λ é indeterminado, podemos determiná-lo agora impondo que o coeficiente de dz na expressão anterior seja nulo, ou seja, que

$$\frac{\partial f}{\partial z} + \lambda\frac{\partial g}{\partial z} = 0 \tag{11}$$

Com isso, temos agora um df independente de z, e podemos localizar seus pontos de máximo impondo que $df = 0$, ou, mais precisamente, que $df + \lambda dg = 0$. Mas isso dá as condições

$$\frac{\partial f}{\partial x} + \lambda\frac{\partial g}{\partial x} = 0 \tag{12}$$

$$\frac{\partial f}{\partial y} + \lambda\frac{\partial g}{\partial y} = 0 \tag{13}$$

Como, adicionalmente, temos a condição dada pela Eq.(11), notamos que o conjunto das equações que determinam os pontos de máximo (bem como o valor de λ) é obtido da seguinte maneira: igualem-se a zero as derivadas parciais da função

$$f + \lambda g \tag{14}$$

A generalização é imediata. Seja $f(x, y, z, u, v)$ a função cujos pontos de máximo queremos localizar, e sejam $g(x, y, z, u, v) = 0$ e $h(x, y, z, u, v) = 0$ condições subsidiárias. Então igualam-se a zero as derivadas parciais da função

$$f + \lambda_1 g + \lambda_2 h \tag{15}$$

onde λ_1 e λ_2 são coeficientes a determinar. Se houver n condições subsidiárias $g_i = 0$, igualem-se a zero as derivadas parciais da função

$$f + \sum_i \lambda_i g_i \tag{16}$$

Os λ_i são denominados *multiplicadores de Lagrange.*

2.3 Exemplo simples

Voltemos ao problema do paralelepípedo e vamos resolvê-lo pelo método de Lagrange. A função cujos máximos procuramos é $f(x, y, z) = xyz$; o vínculo é $g(x, y, z) = xy + yz + xz - K = 0$. Logo, temos de igualar a zero as derivadas parciais da função $f + \lambda g$. Um cálculo simples leva a

$$\frac{\partial f}{\partial x} + \lambda \frac{\partial g}{\partial x} = yz + \lambda(y + z) = 0 \tag{17}$$

$$\frac{\partial f}{\partial y} + \lambda \frac{\partial g}{\partial y} = xz + \lambda(x + z) = 0 \tag{18}$$

$$\frac{\partial f}{\partial z} + \lambda \frac{\partial g}{\partial z} = xy + \lambda(x + y) = 0 \tag{19}$$

Das duas primeiras temos

$$yz + \lambda(y + z) = xz + \lambda(x + z)$$

de onde segue imediatamente que $x = y$. Das duas últimas, analogamente, segue que $y = z$. Logo, $x = y = z$, e se trata de um cubo. Note-se que não foi sequer necessário calcular λ. Assim, mesmo neste caso muito simples, é vantajoso usar o método dos multiplicadores de Lagrange.

3 Aplicações à física

Como primeira aplicação, vamos determinar que movimentos um corpo (sistema macroscópico) pode executar no equilíbrio termodinâmico, ou seja, sob a condição de que a entropia seja um máximo. Seguimos o tratamento do problema dado por Landau [5]. Dividimos o corpo em um grande número de partes pequenas (mas ainda macroscópicas), e sejam M_i, E_i e $\vec{P}_i$ a massa, a energia e o momento da i-ésima parte. A entropia de cada parte é uma função de sua energia interna, isto é, da diferença entre sua energia total E_i e sua energia cinética, dada por $\frac{\vec{P}_i^2}{2M_i}$. A entropia total do corpo, portanto, pode ser escrita

$$S = \sum_i S_i \left(E_i - \frac{P_i^2}{2M_i} \right) \tag{20}$$

Supondo que o corpo seja um sistema isolado, o momento e o momento angular devem ser conservados. Temos, portanto, dois vínculos:

$$\sum_i \vec{P}_i = constante \tag{21}$$

$$\sum_i \vec{r}_i \times \vec{P}_i = constante \tag{22}$$

onde $\vec{r}_i$ é o vetor de posição da i-ésima parte do corpo. As condições de vínculo são duas condições vetoriais, portanto, na realidade, seis condições escalares. Logo, são necessários seis multiplicadores de Lagrange. Mas podemos reuní-los em dois vetores, denotados por $\vec{a}$ e $\vec{b}$. Para determinar o máximo da entropia na presença desses dois vínculos vetoriais, devemos, então, igualar a zero as derivadas parciais da função

$$S_i' = \sum_i \left(S_i + \vec{a}.\vec{P}_i + \vec{b}.(\vec{r}_i \times \vec{P}_i \right) \tag{23}$$

onde $\vec{a}$ e $\vec{b}$ são dois vetores constantes a determinar. A entropia deve ser um máximo em função das energias internas. Derivar em relação à energia interna é equivalente a derivar em relação aos momentos, logo, para o nosso problema, as variáveis são as componentes P_i do momento. Por isso devemos ter

$$\frac{\partial}{\partial P_j} S_i' = 0 \tag{24}$$

ou, mais sucintamente,

$$\vec{\nabla}_{\vec{P}_i} S_j' = 0 \tag{25}$$

onde $\vec{\nabla}_{\vec{P}_i}$ é o gradiente na variável $\vec{P}_i$, ou seja, o operador vetorial

$$\frac{\partial}{\partial P_{ix}} \vec{i} + \frac{\partial}{\partial P_{iy}} \vec{j} + \frac{\partial}{\partial P_{iz}} \vec{k}$$

denotando por $\vec{v}_i = \frac{\vec{P}_i}{M_i}$ a velocidade da i-ésima parte do corpo, temos (veja o cálculo detalhado no apêndice)

$$\vec{\nabla}_{\vec{P}_i} S_j \left(E_j - \frac{P_j^2}{2M_j} \right) = -\frac{\vec{P}_i}{M_i T} = -\frac{\vec{v}_i}{T} \tag{26}$$

onde T é a temperatura absoluta, e usamos o fato de que

$$\frac{\partial S}{\partial E} = \frac{1}{T}$$

considerando-se a entropia como função da energia interna e do volume. O cálculo das derivadas parciais da função S', na Eq.(23), pode agora ser feito sem maiores dificuldades. Obtém-se

$$\frac{-\vec{v}_i}{T} + \vec{a} + \vec{b} \times \vec{r}_i = 0 \tag{27}$$

onde usamos $\vec{\nabla}_{\vec{P}}\vec{a}.\vec{P} = \vec{a}$ e que $\vec{\nabla}_{\vec{P}}\vec{b}.(\vec{r} \times \vec{P}) = \vec{b} \times \vec{r}$. Introduzindo as notações $\vec{u} = T\vec{a}$ e $\vec{\Omega} = T\vec{b}$, temos então:

$$\vec{v}_i = \vec{u} + \vec{\Omega} \times \vec{r}_i \ . \tag{28}$$

O significado físico é claro: o movimento das partes do corpo que é compatível com o equilíbrio é composto de uma translação do corpo como um todo, e de uma rotação, como um todo, com velocidade angular $\vec{\Omega}$.

Por que, então, as massas fluidas em torno de nós não estão se movimentando apenas dessa maneira? Por que não são um sistema isolado: tanto o bombeamento de energia do Sol quanto a ação da gravidade da Terra são ações externas, e, nessas condições, a entropia não tem a obrigação de estar num máximo. Enquanto é fácil blindar o sistema contra a luz do Sol, isto é muito mais difícil no caso do campo gravitacional. A ação combinada do Sol e da gravidade da Terra produz toda a sorte de movimentos "proibidos", como as correntes de convecção nos mares e ares (estas últimas popularmente conhecidas como vento...). Uma aplicação interessante do resultado que obtivemos é o crescimento de cristais em estações espaciais. Como elas estão em queda livre, o campo gravitacional é eliminado; sendo o ambiente fechado, a radiação solar também não está presente. Assim, se está, nessas estações, nas condições de validade do nosso "teorema": como só há rotações e translações, que são fáceis de evitar, consegue-se crescer cristais em condições de total ausência de convecção, o que permite obter cristais perfeitos (e enormes).

3.1 Outro exemplo : correntes elétricas estacionárias

Seja dado um condutor ôhmico, ou seja, tal que a relação entre a corrente e o campo elétrico seja

$$\vec{j} = \sigma\vec{E} \tag{29}$$

onde σ é a condutividade do material. Sabemos que a potência dissipada no sistema é dada por

$$\int dV\vec{j}.\vec{E} \tag{30}$$

Há um teorema, atribuído a Lord Kelvin, que diz que a distribuição de correntes estacionárias em um condutor é aquela que minimiza a potência dissipada. Vamos obter este resultado com os métodos que desenvolvemos até aqui.

Uma distribuição de correntes se diz estacionária se

$$div\ \vec{j} = 0 \tag{31}$$

Para determinar completamente $\vec{j}$ é suficiente, então, dar o valor de $rot\ \vec{j}$. A potência dissipada é dada pela expressão

$$\int dV \vec{j}.\vec{E} \tag{32}$$

O vínculo, neste caso, é a condição de estacionaridade, que é

$$div \vec{j} = 0 \tag{33}$$

O que nós queremos saber é qual função $\vec{j}$ minimiza a potência. Trata-se, portanto, de uma generalização do que fizemos até aqui (que era, dada uma função, saber em que ponto era máxima). O problema se resolve assim: a corrente que gera um mínimo de potência é aquela particular corrente $\vec{j}$ tal que, acrescentando-se uma pequena correção $\delta\vec{j}$ a ela, a potência não muda de valor em primeira ordem. Trata-se de uma generalização imediata do caso anterior. De fato, o ponto em que uma função tem um mínimo (ou máximo) goza da seguinte propriedade: mudando-se ligeiramente essa posição, o valor da função não se altera em primeira ordem (é este o significado de a derivada se anular).Então, exigir que a "variação" da função seja zero em primeira ordem, é equivalente a exigir que as derivadas primeiras se anulem. Seja P a potência dissipada. Temos

$$P = \int dV \vec{j}.\vec{E} \tag{34}$$

Substituímos agora $\vec{j}$ por $\vec{j} + \delta\vec{j}$, onde $\delta\vec{j}$ é um infinitésimo. Obtemos uma nova potência

$$P + \delta P = \int dV(\vec{j} + \delta\vec{j}).\vec{E} \tag{35}$$

A variação de P é $P + \delta P - P = \delta P$, com

$$\delta P = \int dV \delta\vec{j}.\vec{E} \tag{36}$$

Para a corrente que minimiza a potência, devemos ter

$$\delta P = \int dV \delta\vec{j}.\vec{E} = 0 \tag{37}$$

É claro que a corrente que minimiza a potência é a corrente zero. Mas o que nós queremos é saber qual a corrente não-nula e estacionária que minimiza a potência. Temos então que impor o vínculo $div\ \vec{j} = 0$. A expressão cuja variação em primeira ordem deve ser zero é

$$\int dV \left\{ \vec{j}.\vec{E} + \lambda div\ \vec{j} \right\} \tag{38}$$

onde λ é um multiplicador de Lagrange. Porém, neste caso, o multiplicador de Lagrange deve ser uma função $\lambda(x, y, z)$ em vez de um número, como no caso anterior. Acrescentando-se um $\delta\vec{j}$ à corrente e usando $\vec{j} = \sigma\vec{E}$, temos :

$$\int dV \left(\frac{(\vec{j} + \delta\vec{j})^2}{\sigma} + \lambda div(\vec{j} + \delta\vec{j}) \right) \tag{39}$$

cuja variação em primeira ordem, que devemos igualar a zero, é:

$$\int dV \left(2\frac{\vec{j}.\delta\vec{j}}{\sigma} + \lambda div(\delta\vec{j}) \right) = 0 \tag{40}$$

Resta colocar o último termo da integral numa forma mais reveladora. Temos

$$\lambda div(\delta\vec{j}) = div(\lambda\delta\vec{j}) - \vec{\nabla}\lambda.\delta\vec{j}$$

Logo, a integral (40) pode ser escrita

$$\int dV(2\frac{\vec{j}}{\sigma} + \vec{\nabla}\lambda).\delta\vec{j} = 0 \tag{41}$$

onde a parte do integrando que contém um divergente foi desprezada por razões bem conhecidas ($\delta\vec{j}$ se anula no infinito). Como $\delta\vec{j}$ é arbitrário, devemos ter

$$\vec{j} = -\frac{\sigma}{2}\vec{\nabla}\lambda \tag{42}$$

ou

$$rot\ \vec{j} = 0 \tag{43}$$

que completa a determinação da corrente. Como estas são as equaçõeses para correntes estacionárias ôhmicas usuais, o teorema está provado: as correntes ôhmicas estacionárias que minimizam a potência dissipada são aquelas para as quais $rot\ \vec{j} = 0$ (em última análise, são aquelas para as quais $V = Ri$).

References

[1] R. Courant, *Differential and Integral Calculus*

[2] T. Apostol, *Calculus*

[3] E. l. Lima, *Anaálise Matemática*, Vol.2

[4] C. Jordan, *Cours d'Analyse à L'Ecole Polytéchnique* Gauthier-Villars, Paris.

[5] L.D. Landau, E.M. Lifshitz, *Statistical Physics*, Part 1, Pergamon, Oxford.