Multiplicadores de Lagrange e Dualidade em Programação Linear

Uma aplicação de derivadas de funções de uma ou mais variáveis é a determinação de valores máximos e/ou mínimos.

Máximos e Mínimos Irrestritos

A definição de extremos relativos para funções de duas variáveis de entrada é idêntica àquela para funções de uma variável, só precisamos lembrar que estamos trabalhando com funções de duas variáveis (ou mais). Portanto, para fins de completude, aqui está a definição de mínimos relativos e máximos relativos para funções de duas variáveis.

- Uma função $f(x,y)$ tem um mínimo relativo no ponto (a,b) se $f(x,y) \geq f(a,b)$ para todos os pontos (x,y) em alguma vizinhança de (a,b).
- Uma função $f(x,y)$ tem um máximo relativo no ponto (a,b) se $f(x,y) \leq f(a,b)$ para todos os pontos (x,y) em alguma vizinhança de (a,b).

Teorema (Regra de Fermat): Se f tem um máximo ou mínimo local em $(a,b) \in D(f)$ então

$$\begin{cases} \dfrac{\partial f}{\partial x}(a,b) = 0 \\ \dfrac{\partial f}{\partial y}(a,b) = 0 \end{cases}$$

Máximos e Mínimos Restritos – Multiplicadores de Lagrange

A técnica de **multiplicadores de Lagrange** permite encontrar o máximo ou mínimo de uma função $f(x,y)$ quando há restrições nos valores das variáveis de entrada, ou seja se aplica a restrições do tipo $g(x,y) = k$. A restrição g é uma função com o mesmo domínio de f, e k é uma constante.

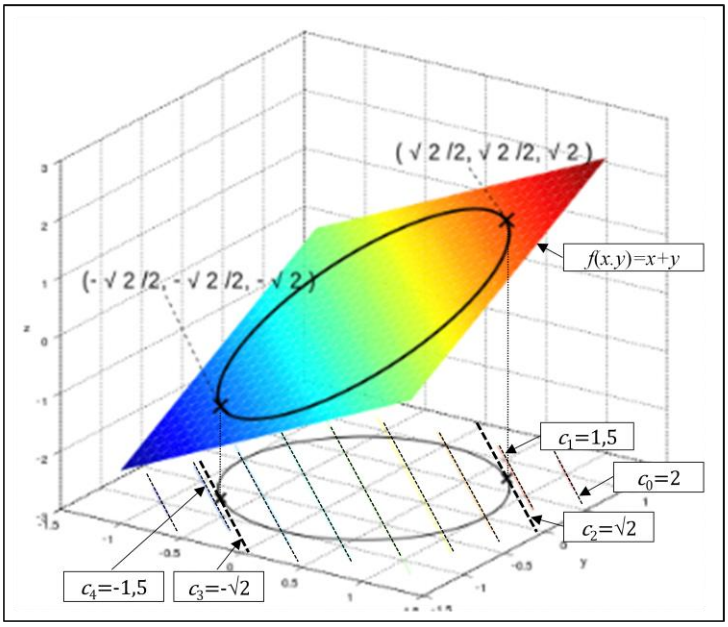
Teorema: Sejam f e g funções que têm derivadas parciais de 1ª ordem contínuas num domínio aberto comum $D \subseteq \mathbb{R}^2$ (D é definido pela restrição g). Se $(x_0, y_0) \in D$ é um ponto extremo de f em D e $\nabla g(x_0, y_0) \neq 0$, então existe um número $\lambda \in \mathbb{R}$, tal que $\nabla f(x_0, y_0) = \lambda \nabla g(x_0, y_0)$.

Este teorema nos dá uma condição necessária (mas não suficiente) para que um ponto $(x_0, y_0) \in D$ seja um ponto extremo de f.

Exemplo 1: Consideramos o caso bidimensional, $z = f(x,y)$ com restrição $g(x,y) = k$. Para facilitar o entendimento suponhamos, sem perda de generalidade, que $g(x,y) = k$ é uma circunferência, e portanto, o domínio de f são os pares de pontos que formam esta circunferência.[a]

[a] https://medium.com/towards-data-science/lagrange-multipliers-kkt-conditions-duality-intuitively-explained-de09f645b068

Consideremos $f(x,y) = z = x + y$ e a restrição $g(x,y) = x^2 + y^2 = 1$. Sejam algumas curvas de nível de $z = f(x,y) = c_i$ com valores explícitos c_i ("c_i chutados"). Geometricamente estamos procurando por pontos da circunferência onde z assume valores máximos ou mínimos.

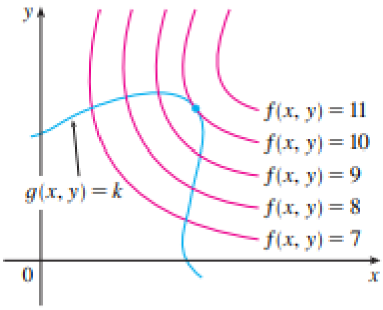


Comecemos na curva de nível com $c_0 = 2$, que não tem pontos na circunferência. Então, claramente o valor máximo de z na circunferência de restrição é menor que 2. Movemos a curva de nível até tocar a circunferência que ocorre quando $c_2 = \sqrt{2}$. Denominemos o ponto do primeiro toque de x^*. É claro que x^* dá um máximo local para z em $g(x,y) = k$, porque se você se afastar de x^* em qualquer direção na circunferência você estará em uma curva de nível com valor c_i menor.

Como a circunferência é uma curva de nível para g, sabemos que ∇g é perpendicular a ela. Sabemos também que ∇f é perpendicular à curva de nível $c_2 = \sqrt{2}$. Uma vez que as duas curvas são tangentes em x^*, estes dois gradientes são paralelos em x^*.

Da mesma forma, se continuarmos descendo as curvas de nível, a última a tocar a circunferência dará um mínimo local e o mesmo argumento pode ser aplicado.

Exemplo 2: $z = f(x,y)$ e $g(x,y) = k$

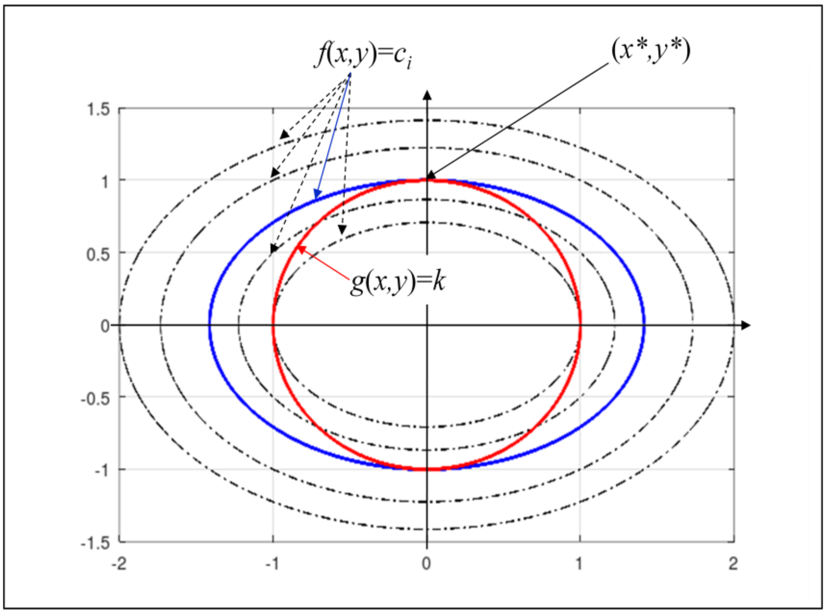


A Figura mostra a curva de $g(x,y) = k$ junto de curvas de nível de f. Estas têm as equações $f(x,y) = c_i$ onde $c_i \in \{7,8,9,10,11\}$. Para maximizar $f(x,y)$ sujeita a $g(x,y) = k$ é preciso determinar o maior valor de c_i, tal que a curva de nível $f(x,y) = c_i$ intercepte $g(x,y) = k$. Parece, da Figura, que isso acontece quando essas curvas se tocam, ou seja, quando essas curvas têm uma reta tangente comum. (Caso contrário, poderíamos aumentar o valor de c_i.) Isso significa que as retas normais ao ponto (x_0, y_0) onde as duas curvas se tocam devem ser as mesmas. Logo, os vetores gradientes são paralelos, ou seja, $\nabla f(x_0, y_0) = \lambda \nabla g(x_0, y_0)$ para algum escalar ∇.[b]

[b] Stewart, James Cálculo, volume 2 / James Stewart ; tradução EZ2 Translate. -- São Paulo : Cengage Learning, 2013.

Então, a principal ideia é procurar pontos onde as curvas de nível de f e a curva g sejam tangentes entre si. Isto é o mesmo que encontrar pontos onde os vetores gradientes de f e g são paralelos entre si. Quando as curvas de nível de duas funções f e g são tangentes num ponto, seus gradientes são paralelos neste ponto.

Exemplo 3: $f(x,y) = x^2 + 2y^2$ no círculo $x^2 + y^2 = 1$ $(g(x,y) = k)$



Função Lagrangeana considerando uma restrição de igualdade

O fato das curvas $f(x,y) = c_i$ e $g(x,y) = k$ serem tangentes não diz nada sobre a magnitude de cada um dos gradientes mas nos interessa a direção. Quando dois vetores apontam na mesma direção significa que podemos multiplicar um por algum escalar positivo para obter o outro. Especificamente, seja (x_0, y_0) um ponto em que as curvas de f e g são tangentes. Como essa tangência significa que seus gradientes estão alinhados/paralelos então podemos escrever

$$\nabla f(x_0, y_0) = \lambda \nabla g(x_0, y_0), \lambda \in \mathbb{R}$$

Nos anos 1700 Joseph Louis **Lagrange** escreveu uma função especial que engloba f e g e uma nova variável λ.

$$\mathcal{L}(x, y, \lambda) = f(x,y) + \lambda(g(x,y) - k)$$

A função $\mathcal{L}$ é denominada de **Função Lagrangeana** ou simplesmente **Lagrangeana**, e a variável λ de **multiplicador de Lagrange**.

As três condições que precisamos resolver para encontrar x, y, e λ passam por todas as derivadas parciais de $\mathcal{L}$ igualadas a 0, ou seja, $\nabla \mathcal{L} = 0$:

$$\begin{cases} \dfrac{\partial f}{\partial x}(x,y) - \lambda \dfrac{\partial g}{\partial x}(x,y) = 0 \\ \dfrac{\partial f}{\partial y}(x,y) - \lambda \dfrac{\partial g}{\partial y}(x,y) = 0 \\ g(x,y) - c = 0 \end{cases}$$

Função Lagrangeana considerando m restrições de igualdade

Seja $f(x) = f(x_1, \ldots, x_n)$ e sejam m restrições $g_i(x) = g_i(x_1, \ldots, x_n) = 0$. O problema é

maximize (ou minimize) $f(x)$, sujeito a $g_i(x) = 0, i = 1, \ldots, m$, com $x = (x_1, \ldots, x_n)^T$

Para investigar máximo(s)/mínimo(s) de $f(x)$ sujeito às m restrições $g_i(x) = 0$, definiremos a Lagrangeana $\mathcal{L}$:

$$\mathcal{L}(x, \lambda_1, \ldots, \lambda_m) = f(x) + \sum_{i=1}^{m} \lambda_i g_i(x)$$

onde λ_i são os multiplicadores de lagrange λ_i.

A condição necessária é

$$\begin{cases} \dfrac{\partial \mathcal{L}}{\partial x_j} = \dfrac{\partial f}{\partial x_j}(x) + \displaystyle\sum_{j=1}^{m} \lambda_i \dfrac{\partial g_i}{\partial x_j}(x), j = 1, \ldots, n \\ \dfrac{\partial \mathcal{L}}{\partial \lambda_i} = g_i(x), i = 1, \ldots, m \end{cases}$$

formada por $(n + m)$ equações em função de $(n + m)$ incógnitas. A resolução do sistema fornece os possíveis candidatos para máximos e mínimos locais.

Função Lagrangeana com uma restrições de desigualdade

Fonte 1: https://people.eecs.berkeley.edu/~klein/papers/lagrange-multipliers.pdf
Fonte 2: http://repositorio.unicamp.br/jspui/bitstream/REPOSIP/306532/1/Padua_SuzanGrazielleBenettide_M.pdf

O método do multiplicador de Lagrange também cobre o caso de restrições de desigualdade do tipo $g_i(x_1, \ldots, x_n) \leq 0$. Ou seja, também podemos resolver problemas do tipo

máximo de $f(\mathrm{x})$, sujeito às condições $g_i(x) \leq 0$, com $x = (x_1, \ldots, x_n)^T \in D$

Caso x^* seja solução do problema de otimização, isto é, x^* é o ponto que maximiza f no conjunto viável D, temos duas possibilidades:[c]

(i) x^* está na fronteira do conjunto viável D, $g_i(x^*) = 0$ (Figura (b)); ou

(ii) x^* está no interior do conjunto viável D, $g_i(x^*) < 0$) (Figura (a)).

[c] https://medium.com/towards-data-science/lagrange-multipliers-kkt-conditions-duality-intuitively-explained-de09f645b068

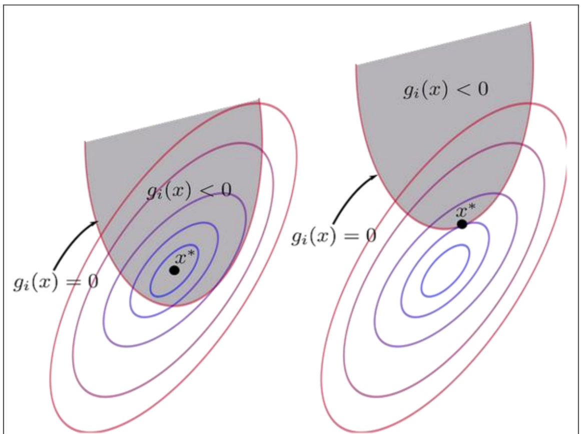


Figura (a) Figura (b)

Detalhando

(i) Se $g_i(x^*) = 0$, dizemos que a restrição g_i está ativa no ponto x^*. Geometricamente, significa dizer que o ponto x^* está na fronteira do conjunto viável D. Então, estamos em um caso parecido com a otimização com restrição de igualdade, valendo a ideia do Teorema dos Multiplicadores de Lagrange de que se x^* é um ponto de máximo de f em D então os gradientes de f e g_i são paralelos nesse ponto. Logo, existe um número real λ^* tal que

$$\nabla f(x^*) = \lambda^* \nabla g_i(x^*)$$

Como o ponto de máximo está na fronteira, e não no interior de D, temos que o gradiente de f no ponto x^* tem que apontar para "fora" da região D, afinal, quando não nulo, ele fornece a direção de maior crescimento da f no ponto. O gradiente de g_i no ponto x^* também fornece a direção de maior crescimento da função g_i no ponto x^*. Como no interior do conjunto viável temos $g_i(x) < 0$ e na fronteira temos $g_i(x) = 0$, então $\nabla g_i(x^*)$ também aponta para fora da região D. Assim, $\nabla g_i(x^*)$ e $\nabla f(x^*)$, além de serem paralelos, têm o mesmo sentido: os dois apontam para "fora"da região viável, o que implica que λ^* é um escalar não negativo

$$\lambda^* \geq 0$$

Desta forma temos

$$\nabla f(x^*) = \lambda^* \nabla g_i(x^*), \lambda^* \geq 0, g_i(x^*) = 0$$

(ii) Se $g_i(x^*) < 0$, dizemos que a restrição g_i não está ativa em x^*. Geometricamente, significa que o ponto x^* está no interior do conjunto viável D. Então estamos em um caso de otimização sem restrição e podemos aplicar **a regra de Fermat**, pois x^* é um ponto crítico de f no interior do conjunto viável D, ou seja, os valores de x independem de g_i, e portanto basta resolver a equação

$$\nabla f(x^*) = 0$$

Desta forma teremos

$$\nabla f(x^*) = 0 \; e \; g_i(x^*) < 0.$$

ou seja, $\lambda^* = 0$.

Podemos unificar as possibilidades (i) e (ii) através da equação $\lambda^*[g_i(x^*) - 0] = 0$, denominada de **condição de complementaridade.** As soluções desta equação são: $\lambda^* = 0$ ou; $g_i(x^*) = 0$.

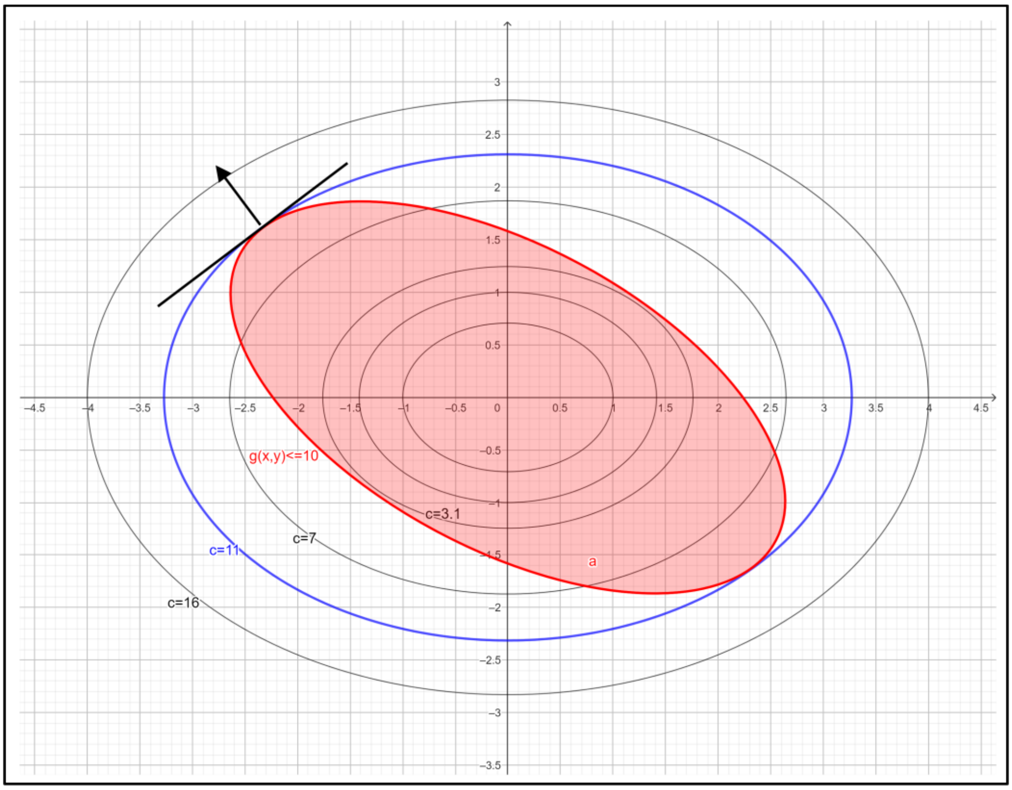
a) Se $\lambda^* = 0$, então $\nabla f(x^*) = \lambda^* \nabla g(x^*) = 0 \times \nabla g_i(x^*) = 0$, isto é, x^* é um ponto crítico de f (podendo ocorrer $g_i(x^*) < 0$ ou $g_i(x^*) = 0$). Se $g_i(x^*) < 0$ então a restrição g_i não está ativa no ponto x^*. Se, porém, $g_i(x^*) = 0$ então que x^* é um ponto crítico de f na fronteira do conjunto viável D.

b) Por outro lado, se $\lambda^* > 0$ temos $g_i(x^*) = 0$, isto é, a restrição g_i está ativa em x^* e, então $\nabla f(x^*) = \lambda^* \nabla g_i(x^*)$, com $\lambda^* > 0$.

Uma conclusão importante é que **se a restrição for do tipo ≤ num problema de maximização, então o multiplicador de Lagrange é não negativo, ou seja, $\boldsymbol{\lambda \geq 0}$.**

Observação importante: Um aspecto útil do método do multiplicador de Lagrange é que os valores dos multiplicadores nos pontos de solução geralmente têm algum significado. Matematicamente, um multiplicador λ_i é o valor da derivada parcial de $\mathcal{L}$ em relação à restrição g_i. Portanto, é a taxa em que poderíamos aumentar a Lagrangeana se elevássemos “a meta” dessa restrição. Mas lembre-se que nos pontos de solução x, $\mathcal{L}(x, \lambda) = f(x)$. Portanto, a taxa de aumento do Lagrangiano em relação a essa restrição é também a taxa de aumento do máximo valor restrito de f com relação a essa restrição. Em economia, quando f é uma função de lucro e g_i é uma restrição de recursos, λ_i é a quantidade (possivelmente negativa!) pela qual o lucro subiria se fosse permitido mais uma unidade do recurso i. Essa taxa é chamada de **preço-sombra** do recurso i, que é interpretada como a quantia que valeria a pena para relaxar a restrição (por P & D, mineração, suborno ou qualquer outro meio).[d]

Exemplo 4: $f(x, y) = x^2 + 2y^2$ na elipse $2x^2 + 3xy + 4y^2 \leq 10$



<u>**Observações**</u>:[e] Os **multiplicadores de Lagrange** são variáveis auxiliares que transformam o problema de otimização restrito em uma forma irrestrita, de modo que o problema se reduz à resolução de um problema de cálculo. Conceitualmente, ele fornece à restrição ou restrições um meio de ‘participar’ do processo de otimização, por meio do qual uma solução ótima deve ser identificada.

[d] http://www.cs.cmu.edu/~ggordon/lp.pdf
[e] https://www.geeksforgeeks.org/lagrange-multipliers/?ref=rp

A **função de Lagrange**, ou apenas **Lagrangiano**, quando queremos maximizar ou minimizar uma função objetivo sujeita a uma ou mais restrições, a Lagrangeana nos ajuda a incorporar essas restrições diretamente no processo de otimização usando multiplicadores de Lagrange.

Geometricamente, a técnica de multiplicadores de Lagrange pode ser vista como o requisito de que o gradiente da função objetivo deve ser paralelo ao gradiente da(s) função(ões) de restrição. Ou seja, em um ponto ótimo, a direção mais íngreme de subida ou descida da função objetivo apontará para uma direção em que a função de restrição não varia. Isso garante evitar violações das restrições ao mesmo tempo em que otimiza uma função objetivo.

<u>Resumo</u>: Os multiplicadores de Lagrange fornecem uma metodologia poderosa para encontrar extremos de uma função sujeita a restrições. Essa abordagem — que combina a transformação de um problema de otimização restrito em um sistema de equações — 'resolve facilmente' soluções ótimas que satisfazem tanto uma função objetivo quanto restrições. O método dos multiplicadores de Lagrange é uma das ferramentas mais úteis, estendendo o cálculo padrão para resolver problemas mais complexos do mundo real, desde modelos econômicos até projetos de engenharia e problemas de física.

Programação Linear e Multiplicadores de Lagrange – Dualidade

Fonte 1: http://www.cs.cmu.edu/~ggordon/lp.pdf
Fonte 2: http://wwwp.fc.unesp.br/~adriana/Pos/PO8.pdf
Fonte 3: https://www.youtube.com/watch?v=u_vNrV14RXM&list=PLjK8Tkmw0e0p8Iv4qfaGniaDRKRkgro_Q&index=4
Fonte 4: http://www.statslab.cam.ac.uk/~rrw1/opt/O.pdf
Fonte 5: https://people.eecs.berkeley.edu/~elghaoui/Teaching/EE227A/lecture7.pdf

Considerações

- $(A^TB)^T = B^TA, \;\; A^T + B^T = (A+B)^T$
- $c \in \mathbb{R}^{n\times 1}, \; x \in \mathbb{R}^{n\times 1}, \; A \in \mathbb{R}^{m\times n}, \; b \in \mathbb{R}^{m\times 1}.$
- ...
- ...

A – <u>PL: minimize $f(x)$, sujeito a $Ax \leq b$ e x livre</u>

Seja o programa linear-pl

$$\begin{array}{ll} \min & f(x) = c^Tx \\ \text{sa} & Ax \leq b \end{array} \quad \Rightarrow \quad \begin{array}{ll} \min & f(x) = c^Tx \\ \text{sa} & Ax - b \leq 0 \end{array}$$

Observe que a variável x é livre. Vamos nos referir ao pl acima como o problema **primal** (P) e à variável de decisão x como a **variável primal**.

Introduzindo os multiplicadores de Lagrange π associados às restrições geramos o **Langrageano**

$$\mathcal{L}(x,\pi) = c^Tx + \pi^T(Ax - b) = -b^T\pi + (A^T\pi + c)^Tx, x \text{ livre}$$

O multiplicador de lagrange, π_i pode ser visto como um parâmetro de penalidade para penalizar a violação de restrição $(A_ix - b_i), i = 1, \dots m$.

Consideremos $X = \{x \in \mathbb{R}^m | Ax \leq b\}$. Então para $\pi \geq 0$, $f(x)$ é limitada abaixo por $\mathcal{L}(x,\pi)$, ou seja

$\forall x \in X, \forall \pi \in \mathbb{R}_+^m : \mathcal{L}(x,\pi) \leq f(x)$.[f]

A **função Lagrange dual**, ou simplesmente a **função dual** é

$$q(\pi) = \min_x \mathcal{L}(x,\pi) = \min_x \{c^T x + \pi^T(Ax-b)\}$$

$$q(\pi) = \min_x \{-b^T\pi + (A^T\pi + c)^T x\} \ = -b^T\pi + \min_x (A^T\pi + c)^T x$$

$$q(\pi) = -b^T\pi + \min_x (A^T\pi + c)^T x$$

Sua solução ótima x depende de π e o mesmo acontece com o objetivo ótimo $q(\pi)$.

A **função dual** fornece uma desigualdade fundamental da teoria da dualidade: limitantes inferiores para os problemas de minimização (superiores para problemas de maximização).

Lema: Para qualquer $\pi \geq 0$, $q(\pi)$ é um limite inferior para $f(x)$ (em particular à solução ótima do pl primal), ou seja, $q(\pi) \leq f(x)$, $\forall \pi \in \mathbb{R}_+^m$ e $\forall x$ tal que $Ax \leq b$.

<u>*Dem*</u>: Sejm $X = \{x : x \in \mathbb{R}^m\}$ e $X_b = \{x : Ax \leq b, x \in X\}$. Então

$$\begin{aligned}
q(\pi) &= \min_{x\in X} \ \mathcal{L}(x,\pi) = \min_{x \in X} \ [f(x) + \pi^T(Ax-b)] \\
&\leq \min_{x\in X_b} \ [f(x) + \pi^T(Ax-b)] \\
&\leq \min_{x\in X_b} \ f(x) \quad (\text{pois } Ax - b \leq 0) \\
&= \begin{array}{ll} \min & f(x) \\ \text{sa} & Ax \leq b \\ & x \text{ livre} \end{array} \\
&= f(x^*), x \in S \ \ c.q.d.
\end{aligned}$$

Corolário:

$$\max_{\pi\geq 0} \ q(\pi) \leq \ f(x^*) \qquad \text{i.e.,} \qquad \max_{\pi \geq 0} \ \min_{x \in X_b} \ \mathcal{L}(x,\pi) \leq \ f(x^*)$$

Da **função dual** concluímos que

$$q(\pi) = \min_x \mathcal{L}(x,\pi) = \begin{cases} -b^T\pi, & A^T\pi + c = 0 \\ -\infty, & \text{caso contrário.} \end{cases}$$

Como estamos interessados em obter limitantes inferiores para $f(x)$, a situação $q(\pi) = -\infty$ não interessa. Ou seja, nos interessa apenas quando $A^T\pi + c = 0$.

Assim, o maior limitante inferior para $f(x)$ é obtido maximizando a **função dual**

$$\max_{\pi\geq 0} q(\pi) = \max_{\pi \geq 0} \left\{-b^T\pi + \min_x (A^T\pi + c)^T x\right\} = \max_{\pi\geq 0} \{-b^T\pi : A^T\pi + c = 0\}$$

Ou seja,

[f] A condição $\pi \geq 0$ garante $\mathcal{L}(x,\pi) \leq f(x)$ pois $(Ax - b) \leq 0$.

$$\begin{array}{ll} \max & -b^T\pi \\ \text{sa} & A^T\pi + c = 0 \\ & \pi \geq 0 \end{array} \quad \text{ou de forma equivalente} \quad \begin{array}{ll} \min & b^T\pi \\ \text{sa} & A^T\pi + c = 0 \\ & \pi \geq 0 \end{array}$$

Este programa linear é denominado de problema **dual** (D) do primal (P) e π de variável dual.

Observações

- Seja o pl primal $\min f(x)$, sujeito a $Ax \leq b, x \in X$.
- A **Lagrangeana** $\mathcal{L}(x,\pi) = f(x) + \pi(Ax - b)$ é uma função definida em X. Por exemplo, se $f(x)$ é definida $x \geq 0$ então NÃO use $x < 0$ em $\mathcal{L}(x,\pi)$.
- Ao definir a Lagrangeana as restrições perdem o efeito. O valor de x não precisa respeitar as restrições.
- Ao minimizar $\mathcal{L}(x,\pi)$ (gerando a **função dual**), π assume um valo fixo e determina-se o minimizador x. A restrição $Ax - b \leq 0$ pode ser violada.
- Por outro lado, se substituir em q o valor de π obtido da resolução do pl dual, vc não consegue calcular o valor de x.

B1 – PL: minimize $f(x)$, sujeito a $Ax = b$ e $x \geq 0$ (pl na forma padrão)

Considere o pl

$$\begin{array}{ll} \min Z = & c^Tx \\ \text{sa} & Ax = b \Rightarrow \\ & x \geq 0 \end{array} \quad \begin{array}{ll} \min Z = & c^Tx \\ \text{sa} & Ax - b = 0 \\ & -x \leq 0 \end{array}$$

Para formar a **Lagrangeana** introduzimos multiplicadores de Lagrange π_i para as m restrições de igualdade e μ_j para as n desigualdades e obtemos

$$\mathcal{L}(x,\pi,\mu) = c^Tx + \pi^T(Ax - b) - \mu^Tx = -b^T\pi + (c + A^T\pi - \mu)^Tx, x \geq 0$$

Se $\mu_j < 0$ para algum j então $\mu^Tx \to -\infty$. Portanto só é garantido a existência de um mínimo finito para $\mathcal{L}(x,\pi,\mu)$ se $\mu_j \geq 0\ \forall j$.

Considerando $\mu \geq 0$, a **função dual** é

$$q(\pi,\mu) = \min_{x\geq 0} \mathcal{L}(x,\pi,\mu) = \min_{x\geq 0}(-b^T\pi + (c + A^T\pi - \mu)^Tx) = -b^T\pi + \min_{x\geq 0}(c + A^T\pi - \mu)^Tx$$

Temos $q(\pi,\mu) = -\infty$ exceto quando $A^T\pi - \mu + c \geq 0$, caso em que é $q(\pi,\mu) = -b^T\pi$, ou seja:

$$q(\pi,\mu) = \begin{cases} -b^T\pi, & c + A^T\pi - \mu \geq 0 \\ -\infty, & \text{caso contrário} \end{cases}$$

O problema **dual** de um pl na forma padrão é maximizar a **função dual** em relação a $\mu \geq 0$, ou seja, Então o pl **dual** é,

$$\begin{array}{ll} \max & -b^T\pi \\ \text{sa} & A^T\pi + c - \mu \geq 0 \\ & \mu \geq 0, \pi \text{ livre} \end{array} \quad \Rightarrow \quad \begin{array}{ll} \max & b^T\pi \\ \text{sa} & A^T\pi + c \geq 0 \\ & \pi \text{ livre} \end{array}$$

B2 – <u>PL: minimize $f(x)$, sujeito a $Ax = b$ e $x \geq 0$ (ignorando $x \geq 0$ na Lagrangeana)</u>

Considere o pl

$$\begin{array}{ll} \min & f(x) = c^T x \\ \text{sa} & Ax = b \\ & x \geq 0 \end{array}$$

Como vimos anteriormente (em **B1**), não é necessário utilizar $x \geq 0$ na formulação do pl dual.

$$\mathcal{L}(x,\pi) = c^T x + \pi^T(b - Ax) = (c^T - \pi^T A)x + \pi^T b, x \geq 0$$

Seja $A = [a_1, a_2, \dots, a_n]$, $c = (c_1, c_2, \dots, c_n)$, $(c^T - \pi^T A) = (c_1 - \pi^T a_1, c_2 - \pi^T a_2, \dots, c_n - \pi^T a_n)$, então

$$\mathcal{L}(x,\pi) = c^T x + \pi^T(b - Ax) = (c_1 - \pi^T a_1)x_1 + (c_2 - \pi^T a_2)x_2 + \cdots + (c_n - \pi^T a_n)x_n + \pi^T b, x \geq 0$$

com domínio $\mathcal{L} = \mathbb{R}_+^n \times \mathbb{R}^m$. A **função dual** é

$$q(\pi) = \min_{x \geq 0} \mathcal{L}(x,\pi) = \min_{x \geq 0} [(c^T - \pi^T A)x + \pi^T b] = \min_{x \geq 0} [(c^T - \pi^T A)x] + \pi^T b$$

$$q(\pi) = \min_{x \geq 0} [(c_1 - \pi^T a_1)x_1 + (c_2 - \pi^T a_2)x_2 + \cdots + (c_n - \pi^T a_n)x_n] + \pi^T b$$

$$q(\pi) = \min_{x_1 \geq 0} (c_1 - \pi^T a_1)x_1 + \min_{x_2 \geq 0} (c_2 - \pi^T a_2)x_2 + \cdots + \min_{x_n \geq 0} (c_n - \pi^T a_n)x_n + \pi^T b$$

Se o coeficiente de alguma variável x for negativo, então $q(\pi) = -\infty$. Se o coeficiente da variável x_j for nulo, então x_j pode assumir qualquer valor não negativo sem alterar a função dual $q(\pi)$. Assim, a primeira parte de $\mathcal{L}(x,\pi)$, que depende de $x_1, \dots, x_n$, é sempre $-\infty$ ou zero. Se for zero, então $\mathcal{L}(x,\pi) = \pi^T b$.

Para cada escolha das variáveis duais, então a **função dual** é facilmente resolvida como uma soma de n subproblemas.

$$q(\pi) = \min_{x_1 \geq 0} (c_1 - \pi^T a_1)x_1 + \min_{x_2 \geq 0} (c_2 - \pi^T a_2)x_2 + \cdots + \min_{x_n \geq 0} (c_n - \pi^T a_n)x_n + b^T\pi$$

Seja o subproblema associado a determinação de x_j, então

$$\min_{x_j \geq 0} (c_j - \pi^T a_j)x_j \Rightarrow \begin{cases} x_j = 0, & \text{se } c_j - \pi^T a_j \geq 0 \\ x_j \to \infty, & \text{se } c_j - \pi^T a_j < 0 \end{cases}$$

Se $c_j - \pi^T a_j < 0$, então $q(\pi) = -\infty$ que não pode ser considerado um limitante inferior. Assim, para determinar um limitante inferior devemos escolher π de tal forma que

$$c_j - \pi^T a_j \geq 0 \Rightarrow q(\pi) = b^T \pi$$

Assim,

$$\pi^T a_1 \leq c_1, \pi^T a_2 \leq c_2, \dots, \pi^T a_n \leq c_n \Leftrightarrow \pi^T A \leq c^T \Leftrightarrow A^T \pi \leq c$$

Então o pl **dual** é,

$$\begin{array}{ll} \max & \pi^T b \\ \text{sa} & A^T \pi \leq c \\ & \pi \text{ livre} \end{array}$$

C – <u>PL: maximize $f(x)$, sujeito a $Ax \leq b$ e $x \geq 0$</u>

Considere o pl

$$\begin{array}{ll} \max & f(x) = c^T x \\ \text{sa} & Ax \leq b \\ & x \geq 0 \end{array} \Rightarrow \begin{array}{ll} \max & f(x) = c^T x \\ \text{sa} & Ax + s = b \\ & x, s \geq 0 \end{array} \Rightarrow \begin{array}{ll} \max & f(x) = c^T x \\ \text{sa} & Ax + s - b = 0 \\ & x, s \geq 0 \end{array}$$

Desconsideremos as restrições $x, s \geq 0$ na construção da **Lagrangeana**, ou seja

$$\mathcal{L}(x, s, \pi) = c^T x - \pi^T (Ax + s - b) = (c^T - \pi^T A)x - \pi^T s + \pi^T b, x, s \geq 0.$$

Podemos encontrar o conjunto Π tal que $\pi \in \Pi$ implica em $q(\pi) = \max_{x,s \geq 0}\{\mathcal{L}(x, s, \pi)\}$ finito, e para $\pi \in \Pi$ calculamos o mínimo de $q(\pi)$.

Considere o termo linear $-\pi^T s$. Se algum elemento $\pi_i < 0$, então podemos tornar $-\pi_i s_i$ grande o quanto quisermos tomando s_i grande. Portanto, existe um máximo finito para $q(\pi)$ se $\pi_i \geq 0 \ \forall i$.

Da mesma forma, o termo $(c^T - \pi^T A)x$ pode ser tornado tão grande quanto quisermos, a menos que $(c_i^T - \pi^T a_i) \leq 0$ para todo i. Assim,

$$\Pi = \{\pi : \pi \geq 0, c^T - \pi^T A \leq 0\} \Rightarrow \Pi = \{\pi : \pi \geq 0, \pi^T A \geq c^T\}.$$

Se escolhermos $\pi \in \Pi$ então

$$\max_{s \geq 0}\{-\pi^T s\} = 0$$

escolhendo $s_i = 0$ se $\pi_i > 0$ e qualquer s_i se $\pi_i = 0$. Similarmente

$$\max_{x \geq 0}(c^T - \pi^T A)x = 0.$$

Portanto, para $\pi \in \Pi$, $q(\pi) = \pi^T b$.

Então o pl **dual** é,

$$
\begin{array}{ll}
\min & \pi^T b \\
\text{sa} & A^T \pi \geq c^T \\
 & \pi \geq 0
\end{array}
$$

A intuição chave por trás da teoria da dualidade em programação linear é a seguinte: Para qualquer problema de otimização convexa, sempre existem configurações das variáveis duais, de modo que o mínimo irrestrito da Lagrangeana em relação às variáveis primais (mantendo as variáveis duais fixas) coincida com a solução do problema original de minimização restrita.

A dualidade surge não apenas de justificativas econômicas, mas também da aplicação de condições de Kuhn-Tucker a problemas de PL ou da função Lagrangeana. (link)

Em suma, a relação entre PL PRIMAL e PL DUAL é dada na tabela

PRIMAL	maximize	minimize	DUAL
restrições	$\leq b_i$ $\geq b_i$ $= b_i$	≥ 0 ≤ 0 livre	variáveis
variáveis	≥ 0 ≤ 0 livre	$\geq c_j$ $\leq c_j$ $= c_j$	restrições

Linear Programming and Extensions, George B. Dantizg (1963), página 125 (link)

Veja, por exemplo, mais em:

- https://www.modelling-energy-systems.org/basics/duality-kkts
- https://loja.grupoa.com.br/otimizacao-continua-asp-teoricos-e-computaciona9788522115013-p1043312?srsltid=AfmBOootLxPfZE6nSMB0aLmDc3F4oawc5h5fxvzNLMy2Md16ydZVIVI8
-