**Otimização sem restrições:**

Suponha que a função  de classe  tenha um extremo, máximo ou mínimo ou ponto de sela, em . Vamos transformar o problema de  dimensões em um problema de uma dimensão através da parametrização:  onde  e . Nesse caso a função de uma variável  tem extremo em  para . A condição de primeira ordem para ser ponto de máximo, mínimo ou sela é que . Derivando  obtemos:





Logo, para ser um ponto extremo é preciso que:



Agora, essa igualdade deve ser verdadeira para qualquer . Escolhendo  então  implica que . Fazendo  percebe-se que  para . Em termos vetoriais isso pode ser escrito da forma:

.

**Condições de Segunda Ordem:**

Se o ponto extremo for de máximo então , e se for de mínimo então . Em outras palavras a condição de segunda ordem é dada pelo sinal da segunda derivada:  é ponto de máximo, e  é ponto de mínimo. Calculando a segunda derivada temos:



Ou seja:



A matriz Hessiana é definida como: . Trata-se de uma matriz simétrica pois . Podemos escrever essa derivada na forma de uma multiplicação de matrizes como . Algumas pessoas usam a notação , mas não gostamos dessa notação na física porque é mesma do Laplaciano  que é um operador escalar e não tensorial como o Hessiano.



Nesse caso a condição para o ponto extremo ser de máximo é que , o que significa que a matriz Hessiana deve ser definida negativa. Já para ser de mínimo a condição é que , o que significa que a matriz Hessiana deve ser definida positiva. Se a matriz Hessiana não for nem definida positiva, nem definida negativa, então o ponto é de sela. Análise da definição de matrizes simétricas é apresentada no apêndice xxx. O fato de que a definição da matriz Hessiana define a concavidade da curva é apresentado no apêndice xxx sobre série de Taylor de funções multivariadas.

**Teorema da função envelope:**

Suponha que desejamos achar o extremo de uma função de  que depende de um parâmetro  , ou seja o problema é:

 sem restrições.

As condições de primeira ordem são:

 onde  é o ponto extremo. Vamos definir uma função de  dada por  . Note que se o problema for de máximo então, para um  fixo, , só tocando a curva  quando . Se for de mínimo então . Por isso a função  é uma função envelope para a função . Pergunta é: quanto vale  ?

**Incluir figuras da função envelope!!**



Entretanto  nos extremos, logo chegamos ao resultado do teorema da função envelope:

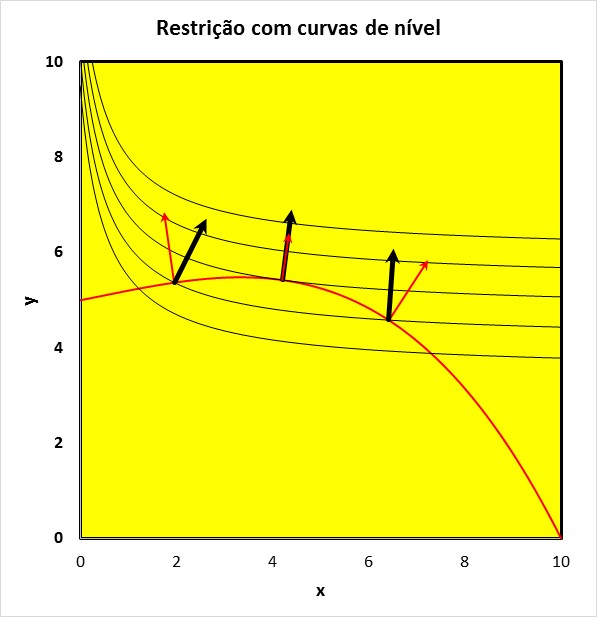
.

**Otimização com restrição de igualdade:**

Suponha agora a seguinte problema: otimizar  sujeito à restrição , ou . Agora nem todo  é permitido, só os que satisfazem à restrição . Existe um raciocínio intuitivo simples que nos permite resolver esse problema. O gradiente de uma função é perpendicular às curvas de nível da mesma. Por um lado sabemos que  porque na curva de nível  é constante. Por outro lado, das regras do cálculo sabemos que:



O que significa que , que nos leva à conclusão de que , com  sendo um deslocamento infinitesimal na curva de nível. A restrição  é uma curva de nível da função . Além disso, como é a própria restrição só deslocamentos sobre a mesma são possíveis. Figura xxx mostra em vermelho a restrição e os gradientes da função . No mesmo gráfico desenhamos em preto diferentes curvas de nível da função  e seus respectivos gradientes nos mesmos 3 pontos da restrição. Note que no ponto de encontro mais à esquerda existe uma componente do gradiente da  ao longo da curva de restrição para à direita. Isso significa um deslocamento para à direita ao longo da curva de restrição aumentará o valor da função . Já no ponto mais à direita a componente do gradiente de  ao longo da curva está na direção esquerda, assim um deslocamento na curva de restrição nessa direção aumenta o valor de . Apenas no ponto em que a restrição tangencia uma curva de nível, com ambos os gradientes paralelos, é impossível aumentar o valor de  através de deslocamentos na restrição.



Esse argumento nos indica, então, que no ponto extremo com restrição os dois gradientes, da restrição e da função  devem ser paralelos. Matematicamente isso é escrito como, ou ainda como . A constante  é chamada multiplicador de Lagrange.

Assim percebe-se que podemos definir uma função Lagrangeana dada por:



Que se comporta como o problema da otimização sem restrições, ou seja:



Assim as condições de primeira ordem para um extremo da função  sujeito à restrição  são que . Essa equação deve ser resolvida em conjunto com a restrição . Entretanto, notamos que  implica em , ou seja, a própria restrição. Dessa forma, usando a função Lagrangeana o nosso problema é resolver o conjunto de equações simultâneas:





Uma demonstração mais rigorosa desse resultado incluindo as condições de segunda ordem é apresentado no apêndice Otimização com Restrições: demonstração rigorosa. Exemplos também são apresentados no apêndice Exemplos de Otimização com Restrição.

**Teorema da função envelope com restrições:**

Suponha que desejamos achar o extremo de uma função de  que depende de um parâmetro , ou seja o problema é:

 com a restrição .

Cujas soluções são dadas por:

 e 

Novamente definimos a função envelope por  e . Só que agora . Mas a restrição continua valendo, ou seja, , logo: . Subtraindo esse zero na derivada de  temos:



Agora sim  no extremo, logo:



.

Esse teorema nos permite extrair uma interpretação do significado do multiplicador de Lagrange  . Note que devido ao fato de que a restrição é nula, , a função objetivo que se deseja otimizar  tem o mesmo valor de . Se derivamos em relação ao parâmetro  vemos que . Ou seja,  é o preço por unidade  da restrição que se paga em não melhorar o objetivo por conta da restrição. Se  poderíamos aumentar a função objetivo diminuindo o valor da restrição , já se  poderíamos aumentar a função objetivo aumentando o valor da restrição , e se  o máximo com restrição coincide com o máximo sem restrição e não há qualquer penalidade devido à restrição.

**Dualidade nos métodos de otimização.**

Suponha o problema de otimização, sem especificar à priori se é de máximo ou de mínimo:

Otimizar  sujeito à restrição .

**Condições de primeira ordem:**

Achamos o Lagrangeano



e procuramos a solução das equações:

 então  logo .

 então .

Resolvendo esse conjunto de equações encontramos ,  e , no qual, obviamente, .

**Condições de segunda ordem:**

O Hessiano orlado é dado pelas derivadas segundas do Lagrangeano

,  e 

com as quais construímos a matriz:

 .

Denotando um subdeterminante por:



Se a solução for um ponto de máximo então:

, , , ou seja, .

Já o ponto for de mínimo então todos os subdeterminantes são negativos e:

.

**Problema DUAL.**

Vamos inverter as funções objetivo e restrição e trocar o problema para:

Otimizar  sujeito à restrição.

Agora o Lagrangeano é dado por:



e encontramos a solução através das equações:

 então  logo .

Comparando com a situação anterior  vemos que basta fazer  para voltar absolutamente às mesmas equações anteriores. A restrição será, obviamente, .



Isso mostra que o mesmo ponto , com ,  , no qual sabemos que , é solução das equações de Lagrange. O problema dual, portanto, fornece a mesma solução do problema original. Resta saber se mantém maximização/minimização ou se troca, maximização se torna minimização e vice-versa. Para isso precisamos das condições de segunda ordem.

Vejamos como fica o Hessiano orlado agora.



Usando o fato de que  re-escrevemos essa derivada segunda como:





As orlas são obtidas através de:

.

Por outro lado  logo  então:





Claro que .

O Hessiano orlado, então, muda para:

 .

No qual o sub-determinante de ordem k é dado por:



Sabemos que multiplicar uma linha ou uma coluna por uma constante multiplica o determinante pela mesma constante, assim podemos pegar a primeira linha e colocar  em evidência, e para cada uma das demais linhas colocar  em evidência. Nesse caso temos que:



Agora colocamos  na primeira coluna para obter:



Ou seja:



Assim mostramos que:



Aqui temos dois casos:

1. **Multiplicador de Lagrange positivo **

Nesse caso o termo não muda o sinal e

.

Se o solução do problema 1 foi de máximo então  e se foi de mínimo então .

O problema dual então será:

 de mínimo quando o original é máximo.

, de máximo quando o original é mínimo.

1. **Multiplicador de Lagrange negativo **

Nesse caso  e  portanto:

,



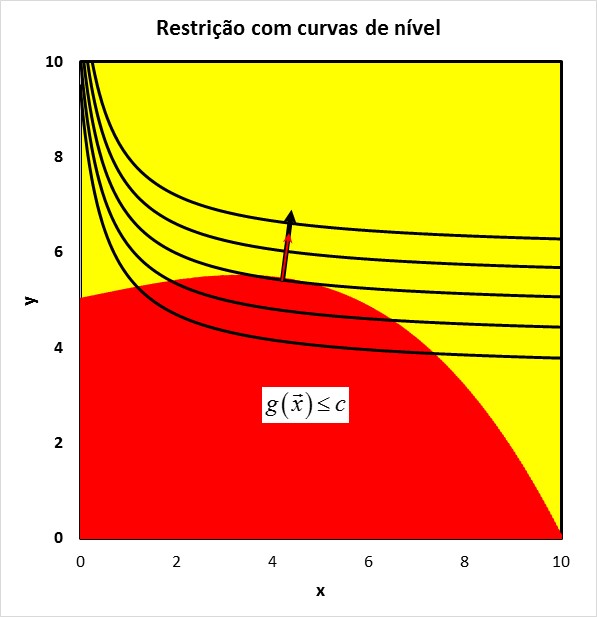
Isso mostra que se o multiplicador de Lagrange é negativo o problema dual tem a mesma característica do que o problema original, ou seja, problema dual de maximização também é de maximização e de minimização também é de minimização.

**Significado do sinal do multiplicador.**

Quando o multiplicador é positivo os gradientes de ambas as funções,  e  apontam na mesma direção, ou seja, não são apenas paralelos. Isso significa que  e  crescem/decrescem na mesma direção. Já se o multiplicador é negativo os gradientes apontam em direções opostas. Se uma das funções cresce em determinada direção a outra decresce. Figura xxx mostra intuitivamente o efeito do sinal de  no problema dual.

**Otimização com restrição de desigualdade:**

Suponha agora a seguinte problema: otimizar  sujeito à restrição  e compará-lo como problema otimizar  sujeito à restrição . Vamos usar a mesma figura



Conside a figura xxx. A região vermelha é a região  , significando que o gradiente da restrição aponta para fora dessa região. Se o gradiente da função também aponta na mesma direção quer dizer que o máximo da função estará na fronteira  e a restrição colou. Por outro lado se o gradiente de está na direção contrária o máximo estará dentro da região  e a restrição não tem qualquer efeito, ou seja, não colou. Nesse caso se procura o máximo sem restrição. No caso da restrição de igualdade, , o ponto extremo tem que estar obrigatoriamente na fronteira, mas não no caso da restrição de desigualdade, . Note então que no problema de achar máximo a restrição só cola se  se  e  estiverem na mesma direção, isto é, . Vale notar que no problema de achar mínimo isso se inverte. Se se  e  estiverem na mesma direção podemos diminuir a função entrando na região de desigualdade. No caso de minimização então a restrição cola se .

Comparando então com o problema de maximização com restrição de igualdade temos que se  a restrição colou e deve-se resolver o sistema de equações:



Por outro lado se a restrição não colou temos um problema de maximização sem restrição e só temos que resolver as equações . Podemos escrever essas duas possibilidades em um conjunto de equações apenas da forma:



Note que agora a condição  leva a apenas duas opções:

1.  então  e a restrição colou, logo:



1.  e  levando ao sistema .

Com esse exemplo podemos extrair as regras:

1. Maximizar  s.r.  então  e devemos resolver as equações: ,  e .
2. Maximizar  s.r.  então  e devemos resolver as equações: ,  e .
3. Minimizar  s.r.  então  e devemos resolver as equações: ,  e .
4. Minimizar  s.r.  então  e devemos resolver as equações: ,  e .

Note que restrições do tipo  podem ser transformadas em , logo o problema pode sempre ser tratado com a restrição do tipo  para maximização e  para minimização bastando redefinir  e . Dessa forma o Lagrangeano sempre será do tipo , sempre exigindo que  e . Essa regra ecvita ter que trocar o sinal de  no Lagrangeano.

**Otimização com restrições mistas:**

1. Maximizar  s.r.  e  então construímos o Lagrangeano dado por:



Devemos resolver as equações:

1.

2.  ou 

3. 

4. 

5. 

1. Minimizar  s.r.  e  então construímos o Lagrangeano dado por:



Devemos resolver as equações:

1.

2.  ou 

3. 

4. 

5. 

As condições de segunda ordem mudam se a restrição cola, usam-se então as regras do Hessiano orlado, ou não cola, usam-se as regras do Hessiano sem orlas.

**Problema de Kuhn-Tucker:**

Kuhn-Tucker querem soluções para variáveis que se podem ser positivas, com um problema do tipo:

Maximizar  s.r.  e .

Usando a técnica anterior o Lagrangeano seria dado por:



As equações conjuntas a serem resolvidas seriam:

1.

2.  ou 

3. 

4. 

5. 

O Lagrangeano de Kuhn-Tucker não inclui as desigualdades , não considear os multiplicadores de Lagrange  , sendo escrito apenas como:



Note que . As equações de Kuhn-Tucker são:



Como  então . Se a restrição  cola escrevemos 

; ;  e .

Equações de Kuhn-Tucker:

1. 
2. 
3. 
4. 
5. 
6. 

**Incluir exemplos de fixação!!**

**Cálculo das Variações:**

Cálculo das Variações começa com o problema da curva brachistocrona levantado por Johann Bernoulli, chamando a atenção de seu irmão Jakob Bernoulli. Leonhard Euler trabalhou sobre o problema e seu livro Elementa Calculi Variationum gerou o termo Cálculo das Variações. Lagrange foi outro nome importante nesse campo e Legendre se preocupou com as condições de segunda ordem para discriminar máximos e mínimos.

Todo o cálculo das variações está baseado na regra de Leibnitz para derivada de integrais demonstrada no apêndice Regra de Leibnitz. Essa regra afirma que:



**Funcional:**

Um funcional associa um número à cada função  de um domínio de funções. O domínio pode restringir os tipos de funções, apenas as diferenciáveis, por exemplo, ou funções que possuem determinados valores em determinados pontos, etc. A expressão abaixo representa um funcional:

onde 

Sabendo  e dada uma trajetória , sabemos calcular , substituir em  e calcular o valor  obtido após a integração. A questão típica do Cálculo das Variações é se existe uma trajetória especial  que torna o valor de  o maior possível, ou o menor possível, ou um extremo?

Vamos começar com e  fixos e restringir o domínio das funções àquelas trajetórias que passam nos pontos  e  com  diferenciáveis. Nesse caso podemos parametrizar  tornando-a função univariada de um parâmetro  da seguinte forma:



Onde  é a trajetório ótima e  um desvio da trajetória ótima. Como  e  são diferenciáveis, então  também é diferenciável. Além disso as restrições dos pontos inicial e final da trajetória implica que:



Assim podemos construir a função univariada da variável :



Que sabemos possuir um extremo em , ou seja,  como condição de primeira ordem. Além disso as condições de segunda ordem são  se o extremo for de máximo e  se o extremo for de mínimo.

Com os limites ,  fixos a derivada da integral facilita, desprezando os termos das derivadas dos limites de integração. Nesse caso:





Agora,  após fazer a integral por partes com  e . O termo uv é nulo porque . Então  implica em:



Em princípio o fato de que uma integral é nula não significa que o integrando seja nulo. Entretanto no nosso caso a integral deve ser nula para qualquer função arbitrária significando que a única forma da integral ser nula sempre é que:



Essa é a famosa equação de Euler-Lagrange.

A equação de Euler-Lagrange representa a condição de primeira ordem para que o funcional seja um extremo. Se é de máximo ou de mínimo será discutido no apêndice Condições de Segunda Ordem do Cálculo Variacional. O problema de otimização do funcional liberando as condições de pontos fixos no início e ou fim da trajetória é discutido no apêndice Condições de Transversalidade.

**Generalizações:**

1. **Mais de uma variável dependente** . Neste caso temos uma equação de Euler-Lagrange para cada variável:



1. **Mais de uma variável independente da forma**:



Neste caso a equação de Euler-Lagrange deve ser estendida para cada t:



1. **Otimização com restrições do tipo 1**. Achar o extremo de:

 sujeito à restrição .

Agora usamos a parametrização para criar duas funções de :





E usamos a técnica dos multiplicadores de Lagrange para encontrar o extreme. Nesse caso:



As soluções são dadas pelas equações  e . Agora:

 e 

portanto:



de onde extraímos que:



Assim percebemos que basta criar o Lagrangeano  com sendo o multiplicador de Lagrange, nest caso, independente do tempo, e resolver a nova equação de Euler-Lagrange:



1. **Otimização com restrições do tipo 2**. Achar o extremo de:

 sujeito à restrição 

Agora como  podemos mudar para o tipo 1 fazendo , e o multiplicador de Lagrange agora pode ser função do tempo. Novamente caímos na equação de Euler-Lagrange na forma:

 e .

**Teoria do Controle Ótimo:**

A teoria do Controle Ótimo é uma matemática do século XX e não mais a matemática do século XIX. O matemático russo L. S. Pontryagin publicou Princípio de Máximo com o qual ganhou o prêmio Lenin em 1962, mesmo ano em que seu trabalho foi traduzido para o inglês. O matemático americano Magnus R. Hestenes da Rand Corporation produziu um report Interno da Rand chamado “A general Problem in the Calculus of Variations with Applications to Paths of Least Time”. Mais tarde ele publicou um paper estendendo os resultados de Pontryagin com o título “On variational theory and Optimal Control Theory”, Journal of SIAM, series A, vol.3 p23-48 (1965), e o livro “Calculus of Variations and Optimal Control Theory”, Wiley&Sons, NY (1966).



L.S. Pontryagin Magnus Hestenes

A teoria do Controle Ótimo está focada em encontrar a melhor trajetória para uma variável de controle , sobre comando do controlador, que leva a variável  a uma trajetória maximizadora/minimizadora de uma determinada função objetivo. A variável de controle  precisa, obviamente, interferir na trajetória de , chamada de variável de estado.

O problema matemático é achar o extremo de:

 sujeito à restrição 

Exemplo: Suponha que exista um estoque  de petróleo e que a taxa de extração, sobre comando do operador do sistema, seja , definindo a relação entre estoque e taxa de extração dada por:



Com extração nula não há consumo e a reserva de petróleo não gera qualquer bem estar. Por outro lado, a extração vai exaurindo a reserva e diminuindo o consumo no futuro. O futuro vale menos do que o presente e a função bem estar acumulada pode ser calculada com uma taxa de desconto da forma:



O problema de otimização, portanto, pode ser colocado na forma: Maximizar



Sujeito às restrições  e . Nesse caso  e a restrição . A variável de controle é  e a de estado .

O problema mais geral do controle ótimo é colocado da seguinte forma: otimizar

 sujeito à restrição  dados  e condições sobre  e .

A variável de controle pode ser descontinua por partes, mas com descontinuidades finitas. Exemplo típico é o ligar e desligar da maioria dos controles de temperatura. Já a variável de estado deve ser contínua, notando que  atua na derivada de , mas não diferenciável porque a derivada pode ser descontínua. A teoria é mais geral do que o cálculo das variações, portanto, pois é capaz de lidar com funções não diferenciáveis. Além disso é capaz de manusear restrições sobre a variável de controle e admite soluções de canto. Por exemplo, suponha o caso on-off em que , só pode assumir valores 0 (desligado) ou 1 (ligado). A trajetória ótima para  pode ser do tipo:



Vamos resolver esse problema com um multiplicador de Lagrange e denominamos:

 variável de controle

 variável de estado

 variável de co-estado [costate variable]

Nosso problema é:

Otimizar:



sujeito às restrições:





 livres.

A restrição  pode ser re-expressa como  logo:



Logo podemos somá-la na função objetivo:



Definindo um Hamiltoniano da forma:



e integrando por partes  re-escrevemos o problema como:



A parametrização é feita da seguinte forma:









 variável de co-estado [costate variable]

A função objetivo univariada é dada por:



Vamos impor :





 usando 

Como  e  são arbitrárias cada um de seus multiplicadores no integrando devem ser nulos independentemente, assim:

 e 

Existe ainda uma equação escondida nesse sistema:

 ou seja .

**Condições de transversalidade:**

Se  é livre, arbitrário, então, 

Se , tempo limite especificado, então 

Se é livre então 

Assim, o problema de achar o extremo de  sujeito às restrições  e , nos leva às seguintes equações de movimento:

Defina o Hamiltoniano  e resolva o conjunto de equações:

 ; ;  e .

Ou seja, sempre se escolhe  que maximiza em todo momento, e resolve-se  ; .

Além disso ainda existe a seguinte propriedade importante na trajetória ótima:





Isso significa que se o Hamiltoniano não depende explicitamente do tempo então  e é uma constante, ou seja, existe uma lei de conservação.

**Cálculo das Variações como um caso particular da Teoria do Controle Ótimo:**

No cálculo das variações o problema de Lagrange era:



Dando origem à equação de Euler-Lagrange:



Vamos re-expressar esse problema no formalismo Hamiltoniano:

Achar o extremo de  sujeito à restrição .

Neste caso  logo:



 por outro lado  portanto  lembrando que  recuperamos a equação de Euler-Lagrange:



Exercícios de fixação da operacionalidade da teoria do controle ótimo estão apresentados no apêndice. Exemplos de fixação da Teoria do Controle Ótimo.

**Apêndice: Série de Taylor**

A Série de Taylor de uma função de classe , i.e., infinitamente diferenciável, pode ser explicada da seguinte forma simples e intuitiva. Desejamos aproximar a função  em torno de  por uma série de potências na forma . O índice  define a ordem da aproximação, com  para aproximação de ordem zero, ou seja, uma reta horizontal, , ou seja uma reta com coeficiente angular , para aproximação de ordem 1,  para a aproximação de ordem 2, e assim por diante. É intuitivo que a melhor aproximação seja aquela em que, no pontotanto a função quanto sua aproximação sejam idênticas. Nesse caso:

, logo .

Para determinar os outros coeficientes  também vamos exigir que os valores das derivadas da função e da aproximação sejam idênticos em , como mostra o gráfico da figura 4, abaixo, para uma aproximação de ordem 2.



Figura 4. Série de Taylor.

Podemos demonstrar que os coeficientes , onde , através dos seguintes passos:

1. 

2. 

3. 

4. 

5. 

6. Impondo que  chegamos a: .

Dessa forma, a Série de Taylor, é dada por:



Que pode ser escrita na forma condensada como:



Um caso particular da série de Taylor é a série de Taylor-McLaurin, para a qual :



Que pode ser escrita na forma condensada como:



Em princípio essa é uma série infinita. Entretanto, uma série infinita só será útil se for possível mostrar que ela converge, ou seja, que pode ser truncada em determinado número de termos e que o erro cometido com essa truncagem tende a zero à medida que o número de termos incluídos cresce. O erro cometido com essa truncagem chama-se Resto. Uma série de Taylor truncada em  só pode ser utilizada para função diferenciável, pelo menos,  vezes. Mas isso relaxa a condição de que a função deve ser infinitamente diferenciável.

**Casos Particulares:**

**Função exponencial: **

A expansão da exponencial é imediata se notarmos que  e que, portanto, . Nesse caso:

.

**Binômio de Newton Generalizado:** 

Lembrando da fórmula do binômio de Newton . Considere a expansão em série de Taylor-McLaurin da função , com  inteiro positivo. Nesse caso, sabemos que , e  e a série de Taylor-McLaurin se torna um polinômio de grau  naturalmente truncada em . Então, vale a igualdade: , que é o próprio binômio de Newton. Os coeficientes dos primeiro e último termos valem 1 pois  e , uma vez que . [[1]](#footnote-1) A série de Taylor não agregou muito valor ao caso das funções de potência em que a expansão binomial de Newton já era muito conhecida.

Entretanto ela pode ser usada para generalizar o binômio de Newton para valores de  negativos ou não inteiros, em que a série se torna infinita. Aí sim, ela agrega grande valor. Se  é negativo não escrevemos  pela dificuldade dos fatoriais de números negativos. Nesse caso é melhor colocar em evidência em cada termo e usar:



**Caso particular do binômio de Newton para:**



Esse resultado pode ser usado para a expansão de . Nesse caso:



Já para  obtemos:

.

**Caso particular do binômio de Newton para:**

Nesse caso teríamos que calcular 



Onde . Nesse caso:



Agora notamos que:



E escrevemos a série como:

.

**Logaritmo:** 

Para o caso do logaritmo, , e , as derivadas podem ser facilmente calculadas usando: , para obter . Desse resultado mostra-se que:



e:

.

**Apêndice: Cálculo do Resto da série de Taylor**

Para calcular o Resto da Série de Taylor, precisaremos do Teorema do Valor Médio de Cauchy. Partimos do teorema de Rolle: se é contínua e diferenciável em  e então existe um número  entre  e em que a primeira derivada dessa função é igual a zero. Pode existir mais de um no intervalo, mas o teorema garante que existe pelo menos um.

**Teorema do Valor Médio de Cauchy (TVM)**

Tomemos duas funções diferenciáveis  e  tal que  e . Com elas construímos uma nova função  e determinamos que obriga a  à satisfazer as condições do Teorema de Rolle, ou seja, que , ou seja, . Resolvendo para  obtemos . Isso significa então que , satisfaz as condições do Teorema de Rolle. Nesse caso, existe um número , i.e., , em que . Por outro lado , logo , ou seja:

.

O teorema do Valor Médio de Cauchy afirma que:

.

No caso particular em que ,  o TVM ele se reduz a .

**Resto da Série de Taylor**

Seja o erro cometido na truncagem da série de Taylor até ordem  e seja . Como  e satisfazem as condições do TVM de Cauchy então:

.

Por outro lado, derivando diretamente a função , obtemos:



Como os termos intermediários se cancelam [soma telescópica]:

.

Derivando diretamente obtemos . Usando esses resultados no TVM de Cauchy:



Percebendo que , então . Logo, . Fazendo  na obtemos:

,

de onde tiramos que existe tal que:

.

O erro em truncar com  termos, também chamado de resto, é dado por  e tende a zero quando  tende a infinito para valores de suficientemente próximos de .

**Apêndice. Série de Taylor de Funções Multivariadas:**

Vamos começar com uma função bivariada  e queremos a expansão em série de Taylor de . Tomando  constante sabemos expandir:





Claro que se tivéssemos começado com  constante teríamos obtido:



Como não pode haver diferença entre as duas formas se percebe que . Agora vamos re-expressar a série da forma:

 onde  é o operador série de Taylor.

Agora queremos a série em ordem crescente em , de tal forma que , então



Por outro lado, lembrando que que as derivadas são operadores lineares, sabemos que o operador



Logo a série de Taylor pode ser expressa como:



O processo pode ser rapidamente generalizado para dimensões maiores do tipo  onde queremos a expansão em série de Taylor de .



Agora fazemos  e multiplicamos e dividimos por  para obter:



Por binômio de Newton sabemos que:



Ou seja:



A série de Taylor em variáveis é dada portanto por:



Até ordem 2 podemos escrever a série de Taylor multivariada como:



Ou ainda como:



A matriz  é chamada de Matriz Hessiana.

**Apêndice. Convexidade de uma curva:**

Função côncava: uma função é côncava se para  onde  é a reta secante que une os pontos  à . Esse é o caso da figura 2(a). Função convexa: uma função é convexa se para  onde .Esse é o caso da figura 2 (b).

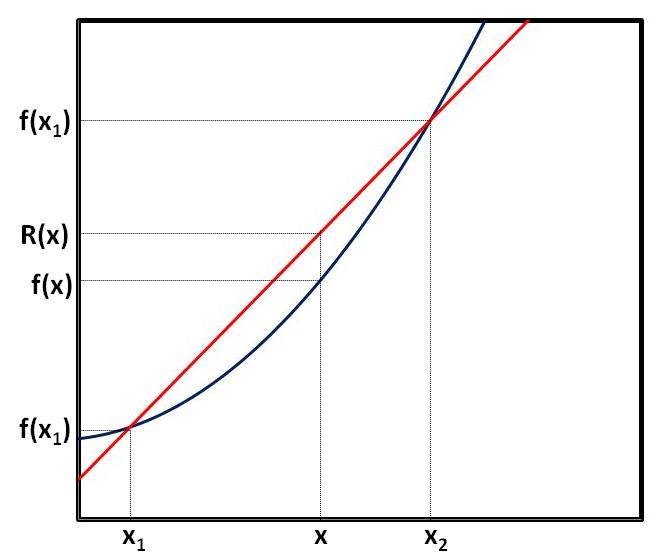
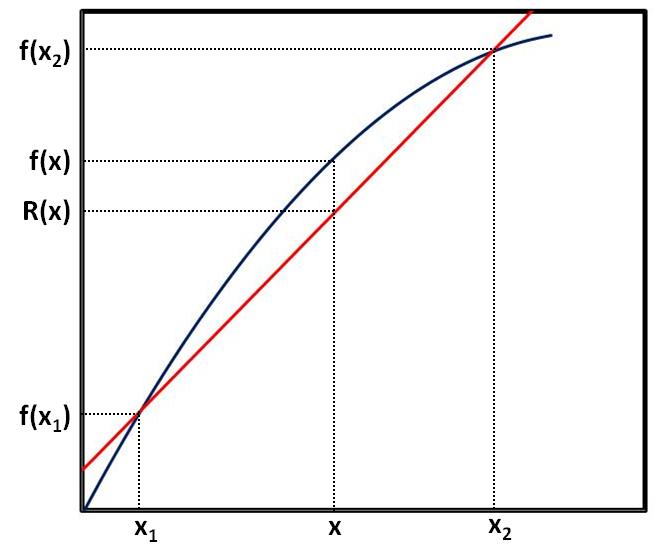


Figura 2. (a) Função côncava. (b) Função convexa.

Vamos construir um  através de um parâmetro  da forma . Se  então  e se  então . Além disso  e , logo pertence ao intervalo . A equação da reta secante que passa pelos pontos e  é dada por , ou seja,  ou, em termos de , , que nos leva, finalmente, a: .

Daí afirmamos que:

1. Se  para então a função  é estritamente côncava
2. Se  para então a função  é estritamente convexa.

Generalização da convexidade para funções de :

Seja um vetor em  e  uma função escalar que associa um número real a um vetor. Então:

1. Se  para então a função  é estritamente côncava.
2. Se  para então a função é estritamente convexa

**Apêndice Otimização com Restrições: demonstração rigorosa.**

Se  então existe uma função inversa local implícita tal que  que pode ser utilizada para eliminar a variável  e caímos em funções de  variáveis dadas por:



e



Note que a  já inclui a restrição, automaticamente satisfeita quando fizemos . Então as condições de ponto extremo da  são as usuais:

 e 

Agora  pois , o que significa que . Por outro lado , então . Como  é uma constante, calculada no ponto  temos que:

 ou ainda que .

Por outro lado  automaticamente. Portanto podemos estender a condição para .

Definindo a função Lagrangeana dada por:  vemos que as condições de primeira ordem para um extremo são que . Essa equação deve ser resolvida em conjunto com a restrição, mas notamos que  implica em , ou seja, a própria restrição. Usando a função Lagrangeana o nosso problema é resolver o conjunto de equações simultâneas:





A condição de segunda ordem é um pouco mais complicada. Vamos usar a notação  para simplificar. Nesse caso temos:



Sabemos que , logo  e . Por outro lado:



Subtraindo as duas equações obtemos:



Agora:



Então:



Definindo um  e usando o fato de que  temos que  , ou seja, , i.e.,  nos garante que está na curva . Nesse caso temos que:



Assim concluímos que:



**Hessiano orlado:**

O Hessiano orlado é definido como: . Note as duas orlas nas primeira linha e primeira coluna da matriz Hessiana. A condições de segunda ordem dos pontos de máximo e mínimo serão dadas pela definição do Hessiano orlado discutido no apêndice **Definição de matrizes simétricas com restrição.** Note que o Hessiano orlado poderia sair diretamente do Lagrangeano se incluirmos  como uma variável pois:

 e 

**Apêndice. Definição de matrizes simétricas:**

Dizemos que a matriz  é definida positiva se  para qualquer . A matriz será definida negativa se  para qualquer .

**Teorema 1:**

Se  é simétrica e definida positiva ou negativa então todos os elementos da diagonal ou são positivos ou são negativos. Prova: como a propriedade é válida para qualquer  basta escolher  o que significa que . Logo se  então  e se  então  para qualquer .

Teorema 2: se  é definida positiva ou negativa então todas as submatrizes  serão definidas positivas ou negativas.

Prova: como a propriedade é válida para qualquer  basta escolher  o que significa que .

Teorema 3.

1. Se  é definida positiva então , ou ainda,  para qualquer .
2. Se  é definida negativa então , ou seja, o sinal do determinante vai alternando na forma .

A melhor forma de prova esse teorema é diagonalizar as submatrizes, mas é instrutivo analisar o caso de uma matriz  e outra  com a técnica de completar quadrado.

Matriz :













 onde .

Se  e como  e  é preciso que  e  o que implica em  e . Por outro se  é preciso que  e  o que implica em  e .

Matriz :

























Então se  exigimos que ,  e  o que implica em ,  e . Mas se  é preciso que ,  e  o que implica em ,  e .

No entanto a melhor forma de provar o caso geral é usar a técnica de diagonalização de matrizes.

Já sabemos que se a matriz  é definida positiva ou negativa então as sub-matrizes  possuem a mesma definição. Sabemos que uma transformação de similaridade do tipo preserva o determinante e queremos descobrir o sinal de . Seja  a matriz que diagonaliza , simétrica, ou seja , então podemos fazer , ou seja, . Agora definimos  logo  logo  .

Mas  então . Logo se  então , mas se se  então  . Se  então , mas se  então . Como a transformação de similaridade preserva os determinantes então, a condição  implica em  e a condição  implica em .

**Apêndice: Definição de matrizes simétricas com restrição.**

Suponha agora que os vetores possíveis devem obedecer a uma restrição do tipo: . Nesse caso as duas variáveis  e  não são mais independentes, mas devem obedecer a relação .

Assim 



E o sinal de q será definido pelo sinal de . Agora percebemos que  logo em lugar de encontrar os sinais de dois subdeterminantes encontramos o sinal de apenas um mas com uma orla.

**Generalizando:**

Sem restrição com uma matriz Hessiana da forma  então a matriz é definida positiva se , ,  até  forem todos positivos. A matriz será definida negativa se  e .

Com restrição:  será definido positivo se os determinantes , , ..., 

**Apêndice. Definição de Matrizes e concavidade das curvas:**

A definição da matriz Hessiana também define a concavidade da função. Se  é estritamente côncava ou convexa então:

 para .

Então também vale para  com  mas com . Nesse caso  então:



Agora, por série de Taylor até segunda ordem sabemos que:



Que pode ser escrita como:



Fazendo  temos que:



Logo:









Como  então  e a concavidade é definida por:



ou seja, se o Hessiano for definido negativo a função é côncava e se for definido positivo a função é convexa. É intuitivo que funções côncavas possuam máximos enquanto funções convexas possuam mínimo.

**Apêndice. Exemplos de Otimização com Restrição:**

1. Maximizar  s.r.  .









Das equações (1) e (2) temos que  logo . Jogando essa relação na restrição [equação (3)] temos 

A matriz Hessiana orlada

.

Entre os sub-determinantes  só o  interessa pois uma restrição começa no 1+1+1=3, e troca-se o sinal em relação ao determinante sem restrição. determinante

.

Como  com uma restrição o  sem restrição seria negativo, matriz definida negativa. Assim ponto é de máximo.

Esse problema também é fácil de resolver por substituição. Da restrição temos que . Substituindo na função objetivo , logo  e  e . A segunda derivada , logo o ponto é de máximo.

1. Maximizar  s.r.  e .













Da equações (1) ou  ou . Mas  entra em contradição com . Logo, obrigatoriamente, .





Extraindoe substituindo  com duas opções  ou . Novamente se  então  em contradição com . Logo  e  e .

Para 

Para 

Em qualquer um dos dois casos  a restrição  , logo ,  e  . Assim temos 4 pontos extremos:



Agora vamos ao Hessiano Orlado:

.

Como  então



São 3 variáveis e 2 restrições então só o sub-determinante total  interessa. Podemos usar a coluna ou a linha 3 cheia de zeros para diminuir a ordem do determinante:



Agora usamos a linha 2 para diminuir mais uma ordem:



 como  então:



Para  portanto  logo se trata de ponto de máximo.

Para  portanto  logo se trata de ponto de mínimo

Assim os pontos  e  são pontos de máximo enquanto os pontos  e  são pontos de mínimo. Voltando à função objetivo vemos que  logo  é o valor máximo. Já  logo  é o valor mínimo.

**Apêndice. Regra de Leibnitz:**

Seja . Queremos calcular .









Agora, seja  tal que , ou seja, , então:



Então:





Logo:



Ou seja, a regra de Leibnitz de derivar integrais é:



Casos particulares?

1. Os limites não dependem de . Nesse caso:



1. O integrando não depende de . Nesse caso:



**Apêndice. Condições de Segunda Ordem do Cálculo Variacional:**

Já sabemos que  onde  e . Aplicando a segunda derivada temos:



Que pode ser escrito na forma matricial:



Assim:

 se a matriz  for definida positiva e o extremo é de mínimo.

 se a matriz  for definida negativa e o extremo é de máximo.

Logo, usando as regras dos determinantes das matrizes definidas positivas e negativas, temos que o extremo será de máximo ou mínimo se . Se essa condição não for satisfeita o ponto é de sela. Será mínimo se  e máximo se . Estas são as condições de Legendre.

**Apêndice. Condições de Transversalidade:**

Suponha agora o problema de encontrar o extremo de  sem a exigência de que  e ou sejam fixos. Para simplificar vamos manter o ponto inicial fixo em . A nova parametrização, considerando  a trajetória ótima,  o tempo final ótimo e  o y final ótimo, será:







Agora o funcional é dado por:



E devemos derivar também o limite de integração usando a regra de Leibnitz, da forma:



Logo



A integral por partes agora também deve ser feita com cuidado porque o termo uv não é mais nulo:



Substituindo no resultado anterior temos:



Agora 



Logo

 e no limite  . Substituindo na equação acima temos:



O extremo será dado pela condição:



Novamente a função  é arbitrária significando que a integral e os outros dois termos devem ser nulos independentemente. Para a integral ser nula continua valendo a equação de Euler-Lagrange:



Mas a ela devemos acrescentar as condições de transversalidade:



**Exemplos de fixação:**

Considere o problema de encontrar a distância à origem em um plano . A trajetória agora é  e manteremos a notação fazendo . Nesse caso  e  . A função objetivo . Nesse caso  e . Note que  o que significa que  sempre, e o problema é de mínimo.

Equação de Euler-Lagrange  nos leva a  ou seja , logo . Assim . Se  então . Como , então  e a reta  que passa na origem é uma solução. Retas gerando distâncias mínimas são esperadas. Falta determinar o coeficiente angular da reta.

Caso 1.  e  , então , logo  e a reta  é a trajetória de menor distância entre a origem e o ponto  .

Caso 2.  mas  , então vamos impor que  . Mas ,  logo  logo  e a reta horizontal  é a trajetória de menor distância entre a origem e o ponto .

Caso 3.  mas  , então vamos impor que  . Mas ,  e  logo . Entretanto, para que  é preciso que e a reta horizontal  e  . A trajetória é a reta vertical entre a origem e o ponto .

Caso 4. O ponto final deve cair na reta . Nesse caso impomos a condição , com  temos que a condição dada por  ,. Mas, logo  e . Nesse caso  é a reta perpendicular à  que passa pela origem. O ponto de encontro das retas é dado por  ou seja , ou seja, .

Exercício: Considere o problema da trajetória de menor distância da origem até o círculo  . Mostre que a trajetória é a reta que passa pela origem e pelo centro do círculo  . Determine  e  .

**Apêndice. Exemplos de fixação da Teoria do Controle Ótimo:**

Exemplos de fixação:

1. Maximizar  s.r.  e  .

Nesse caso  ,

 logo , então:

 logo constante. Como  então , logo . Neste caso  e .

1. Vamos apresentar um problema em que a variável de controle pode ser descontínua. Maximizar  s.r. ; ;  e .

Nesse caso .

Maximizar  em relação à :  . Temos que escolher  que maximiza , o qual é uma função linear de . Assim se  o  que maximiza  será , o maior valor permitido para o . Já se  o  que maximiza  será , o menor valor permitido. Dessa forma podemos escrever:



Notando que o fato de que é uma função linear de implica que  sempre, para qualquer , e a condição  sempre, não pode valer nesse caso para determinar . O critério claro é que escolhendo  que maximiza  o tempo todo teremos  máximo.

Encontrando :

 logo uma equação diferencial de primeira ordem. Usando a solução , temos que , ou seja,  e , ou seja, . Dessa forma .

Por outro lado as condições de transversalidade impõem que  logo  e , portanto, . Para   e, claro, . O tempo limite em que  é definido por , ou seja, . Aplicando o logaritmo de ambos os lados, , ou seja, . Com isso temos:



Falta encontrar :

, logo . Como é constante por partes, podemos tomar a solução  e , logo,  de onde tiramos e , com . Temos duas regiões  e .





Usando a condição inicial  tiramos que  logo . A outra condição é que é contínua, então , logo  de onde extraímos que . Finalmente temos:



1. Maximizar  s.r. ; ;  e .

Nesse caso  e  e .

 logo 



 logo . Com a condição inicial  temos  .

A outra condição é que  então  ou , logo . Só falta encontrar  . Para isso usamos  , então  logo  e . As raízes de  são  , sendo a única real e positiva . Assim determinamos todas as variáveis e trajetórias:

; ;  e .

**Apêndice. Formalismos Hamiltoniano e Lagrangeano da Mecânica Clássica:**

Vamos usar os dois formalismos, do Cálculo das Variações e da Teoria do Controle Ótimo sumarizados na tabela abaixo.

|  |  |
| --- | --- |
| Cálculo das Variações | Teoria do Controle Ótimo |
| Problema: Otimizar  sujeito à restrição . |  |
| Usa a Lagrangeana para transformar o problema em: | Problema: Otimizar  sujeito às restrições . |
| Trajetória ótima segue equação de Euler-Lagrange: | Usa a Hamiltoniana  e as trajetórias ótimas seguem equações de Hamilton-Jacobi:  ; ; ; |

**Mecânica Clássica no Cálculo das Variações.** Trocam-se as leis de Newton pelo princípio da Ação Mínima em que a trajetória é aquela que minimiza  . Agora basta mostrar que que usando:



Recuperamos as leis de Newton.

 e , portanto a equação de Euler-Lagrange  nos leva a:



Vale notar que  é o momento.

Generalizando o Cálculo das Variações para  coordenadas  temos que resolver  equações de Euler-Lagrange do tipo . Se temos um sistema com  partículas com 3 graus de liberdade (x,y,z) para cada partícula teremos  coordenadas no total que vamos denotar por .

A grande vantagem do formalismo Lagrangeano na Mecânica Clássica é a possibilidade de mudar as variáveis para satisfazer restrições específicas. Então vamos mudar das coordenadas cartesianas  para um conjunto de coordenadas dado pelas transformações:

 e sua transformação inversa .

Nas coordenadas cartesianas sabemos que o momento é dado por  então, por analogia, vamos definir um momento generalizado através de . Suponha uma transformação de coordenadas em que não há uma dependência explícita do tempo, então:



A força generalizada é definida por . Se as forças são conservativas então existe uma função potencial  tal que  e  só depende de  mas não de  e , então:

 logo:



Novamente obtemos um Lagrangeano  e as equações de Euler-Lagrange em termos das coordenadas generalizadas  ,  e :



**Formalismo Hamiltoniano:**

Primeiro vamos analisar o caso univariado na  coordenada. No formalismo Lagrangeano achar o extremo de  nos leva às equações de Euler-Lagrange , ou seja,  ou seja . No formalismo Hamiltoniano queremos o extremos de  sujeito à restrição  sobre a qual fazemos a transformação de Legendre:  que nos leva às equações de Hamilton:

, logo  . O significado do mulitplicador de Lagrange nesse caso é que  é o momento generalizado. Além disso:

 portanto 

e

 logo 

As equações de Hamilton-Jacobi são:  e .

O Hamiltoniano é obtido do Lagrangeano através da transformação de Legendre: . Usando os mesmos passos clássicos que levaram do Lagrangeano para o Hamiltoniao pode-se mostrar que .

1. A propriedade dos fatoriais é que  e . Fazendo  temos que  que leva a . Fatoriais de números inteiros negativos divergem pois fazendo temos que  ou . [↑](#footnote-ref-1)