Apostila de Métodos Quantitativos - UERJ
Professor: Pedro Hemsley

Derivada Direcional e Gradiente

A regra da cadeia pode ser utilizada para determinar a taxa de variação de uma função f(x_1,...,x_n) em um ponto x* e em uma direção v = ($v_1, ..., v_n$). Considere a equação paramétrica da reta que passa por x* na direção v:

$x = x^* + tv$

Para determinar como f muda ao longo dessa linha, basta fazer:

$$g(t) = f(x^* + tv) = f(\underset{x_1}{x_1^* + tv_1}, ..., \underset{x_n}{x_n^* + tv_n})$$

Use a regra da cadeia para determinar a derivada em t = 0:

g'(t) = $\frac{df}{dx_1}(x_1^* + tv_1) * \frac{dx_1}{dt} + ... + \frac{df}{dx_n}(x_n^* + tv_n)\ \frac{dx_n}{dt}$
 = $\frac{df}{dx_1}(x_1^* + tv_1)\ v_1 + ... + \frac{df}{dx_n}(x_n^* + tv_n)v_n$
g'(t = 0) = $\frac{df}{dx_1}(x_1^*)\frac{dx_1}{dt} + ... + \frac{df}{dx_n}(x_n^*)\frac{dx_n}{dt}$
 $= \begin{bmatrix} \frac{df}{dx_1} & \cdots & \frac{df}{dx_n} \end{bmatrix} \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix}$ = Df(x) v $\longrightarrow$derivada de f em x^* na direção v.

Exemplo: v = e_1= $\begin{bmatrix} 1 & 0 & \cdots & 1 \end{bmatrix}$: Como f muda ao longo do eixo x_1?

$\begin{bmatrix} \frac{df}{dx_1}(x^*) & \cdots & \frac{df}{dx_n}(x^*) \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \frac{df}{dx_1}(x^*)$: Apenas a derivada parcial em relação a x_1.

Exemplo: Função de Produção: y = $4K^{\frac{3}{4}}L^{\frac{1}{4}}$

Qual o aumento da produção no ponto (10000 , 625) na direção (1 , 1) ? Ou seja, aumento proporcional dos insumos.

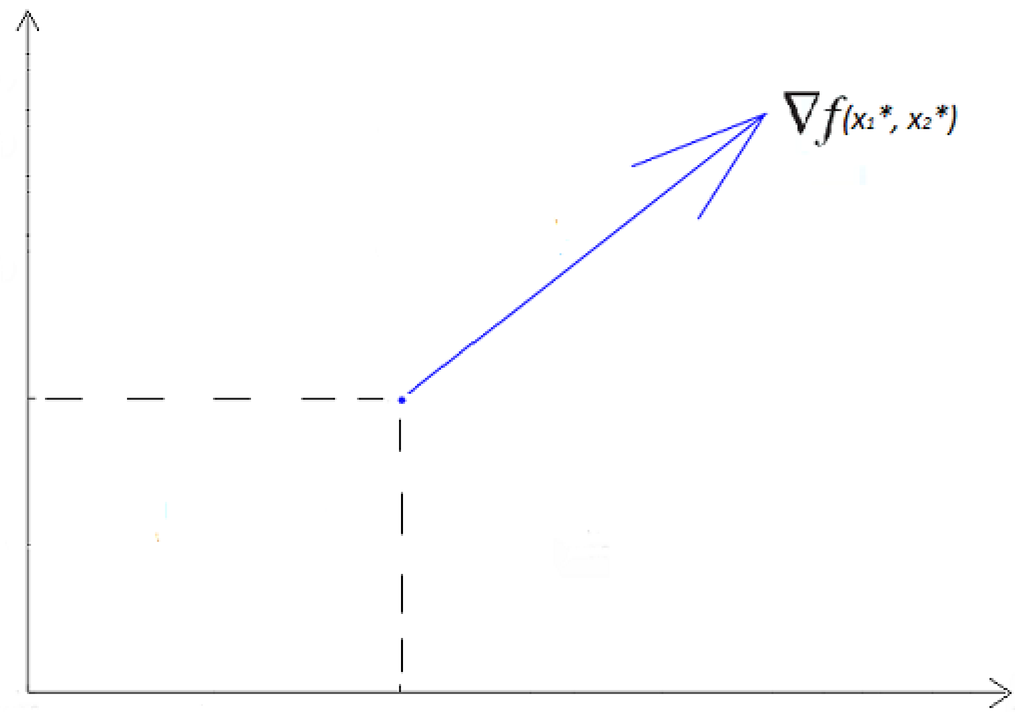
$\frac{dy}{dK}(10000 , 625)\ v_1 + \frac{dy}{dL}(10000 , 625)v_2 = \frac{dy}{dK}(10000, 625)\cdot 1 + \frac{dy}{dL}(10000, 625)\cdot 1$;

Primeiro, calcularemos as derivadas parciais:
$\frac{dy}{dK}$(K , L) = 3 $(\frac{L}{K})^{\frac{1}{4}}$ $\frac{dy}{dL}$(K , L) = $(\frac{K}{L})^{\frac{3}{4}}$

Agora, substituiremos as coordenadas do ponto (10000 , 625):
$\frac{dy}{dK}$(10000 , 625) = 3 $(\frac{625}{10000})^{\frac{1}{4}}$= 1.5 $\frac{dy}{dL}$(10000 , 625) = $(\frac{10000}{625})^{\frac{3}{4}}$ = 8

Fazemos a última substituição na função e acharemos o aumento da produção no ponto (10000 , 625):
$\frac{dy}{dK}(10000 , 625) \cdot 1 + \frac{dy}{dL}(10000 , 625)\cdot 1 = 1.5 \cdot 1 + 8 \cdot 1 = 1.5 + 8 = 9.5$

O gradiente pode ser interpretado como um vetor em R^n a partir do ponto x^*, e é escrito $\bigtriangledown f(x^*)$.



Então: $\bigtriangledown f(x^*)\cdot v = \sum \frac{df}{dx_I}(x^*)$\cdot v$_i$: Produto interno dos vetores $\bigtriangledown f(x^*)$ e v.

Suponha ||v|| = 1. Então:

$\bigtriangledown f(x^*) \cdot v = || \bigtriangledown f(x^*)|| \cdot ||v|| \cdot cos\theta = || \bigtriangledown f(x^*)|| \cdot cos\theta$.

Em que direção a função f(x) aumenta mais rapidamente ?

$v = cos\theta \in [-1, 1]$: Logo, $\bigtriangledown f(x^*) \cdot v = || \bigtriangledown f(x^*)|| \cdot cos\theta$ assumie o maior valor possível quando cos$\theta = 1$. Mas cos$\theta = 1 \Longrightarrow \theta = 0$. Ou seja, $\bigtriangledown f(x)$ e v são PARALELOS na direção de maior crescimento de f..

Então: $\bigtriangledown f(x^*) \cdot v$ assume o maior valor possível (direção de maior crescimento de f) quando $\bigtriangledown f(x)$ e v tem a mesma direção. Logo, $\bigtriangledown f(x*)$ aponta na maior direção de crescimento de f.

Até aqui: f: $R^n \rightarrow R$ na maior parte do tempo. Extensão: f: $R^n \rightarrow R^m$: m variáveis dependentes.

n variáveis exógenas f: $R^n \rightarrow R^m$

m variáveis endógenas

Exemplo: Firma multiproduto: $\left\{ \begin{array}{ccc} q_1 & = & f_1(x_1, ..., x_n) \\ q_2 & = & f_2(x_1, ..., x_n) \\ \vdots & \vdots & \vdots \\ q_m & = & f_m(x_1, ..., x_n) \end{array} \right.$

Ou seja:
f(x) = (f_1(x_1,...,x_n), $\cdots$, f_m(x_1,...,x_n)) : f: $R^n \to R^m$
Ou seja, há m funções do tipo f: $R^n \to R$

Definimos a derivada de f em x* como:

$$\text{Df(x*)} = \begin{bmatrix} \frac{df_1}{dx_1}(x*) & \frac{df_1}{dx_2}(x*) & \cdots & \frac{df_1}{dx_n}(x*) \\ \frac{df_2}{dx_1}(x*) & \frac{df_2}{dx_2}(x*) & \cdots & \frac{df_2}{dx_n}(x*) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{df_m}{dx_1}(x*) & \frac{df_m}{dx_2}(x*) & \cdots & \frac{df_m}{dx_n}(x*) \end{bmatrix}$$

Essa é matriz jacobiana, ou apenas 'jacobiano', e representa a derivada de f. Dessa maneira, temos uma matriz m x n, que por sua vez representa uma função linear f: $R^n \to R^m$: f(x) = A*x, x $\in R^n$, f(x)$\in R^m$, A_{mxn}. Ou seja, o jacobiano é uma aproximação linear para f. Na prática, trabalhamos como se fossem m funções (ou seja, o jacobiano 'empilha' m gradientes).

É possível estudar a regra da cadeia naturalmente.

A derivada de uma função é uma nova função, que também pode ser diferenciável. Quanto a continuidade e diferenciabilidade, podemos classificar as funções como:

c^0: funções contínuas.
c^1: funções diferenciáveis com derivada contínua.
c^2: funções diferenciáveis (2x) com derivada contínua (2x).
$\vdots$
c^k: Análogo.
c^∞: As derivadas sempre existem e são sempre diferenciáveis.

Podemos então definir derivadas cruzadas para k $\geq$2:

f(x_1,...,x_n)
$\frac{df}{dx_i}$(x_1,...,x_n): também é diferenciável
$\frac{d(\frac{df}{dx_i}(x_1,...,x_n))}{dx_j} = \frac{d^2f}{dx_idx_j}(x_1,...,x_n)$
Se i $\neq$j, temos uma derivada cruzada.

Exemplo: $u(x_1,x_2) = x_1^{\frac{1}{2}}x_2^{\frac{1}{3}}$

$\frac{du}{dx_1} = \frac{1}{2}x_1^{-\frac{1}{2}}\ x_2^{\frac{1}{3}}$
$\frac{du}{dx_2} = \frac{1}{3}x_1^{\frac{1}{2}}\ x_2^{-\frac{2}{3}}$

$\frac{d^2u}{dx_1^2} = -\frac{1}{4}x_1^{-\frac{3}{2}}\ x_2^{\frac{1}{3}}$
$\frac{d^2u}{dx_2^2} = -\frac{2}{9}x_1^{\frac{1}{2}}\ x_2^{-\frac{5}{3}}$

$\frac{d^2u}{dx_1dx_2} = \frac{1}{6}x_1^{-\frac{1}{2}}\ x_2^{-\frac{2}{3}}$
$\frac{d^2u}{dx_2dx_1} = \frac{1}{6}x_1^{-\frac{1}{2}}\ x_2^{-\frac{2}{3}}$

Note que $\frac{d^2u}{dx_1dx_2} = \frac{d^2u}{dx_2dx_1}$

Teorema de Young: se y = f(x_1,...,x_n) pertence à classeC^2, então $\frac{d^2u}{dx_i dx_j} = \frac{d^2u}{dx_j dx_i}$

Logo, a ordem da diferenciação não altera o resultado para uma função C^2 qualquer.

As derivadas segundas parciais são organizadas na matriz HESSIANA:

$$D^2f(x)=\begin{bmatrix} \frac{d^2f}{dx_1^2} & \frac{d^2u}{dx_2dx_1} & \cdots & \frac{d^2u}{dx_ndx_1} \\ \frac{d^2f}{dx_1dx_2} & \frac{d^2f}{dx_2^2} & \cdots & \frac{d^2u}{dx_ndx_2} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{d^2f}{dx_1dx_n} & \frac{d^2f}{dx_2dx_n} & \cdots & \frac{d^2f}{dx_m^2} \end{bmatrix}$$

Pelo teorema de Young, essa matriz é simétria para funções C^2.

Objetivo: - Maximizar função-linear
- Conjunto composto
- Solução interior x Solução de canto.

Formas Quadráticas

Extensão de funções quadráticas simples. Função mais simples que admite solução interior. Admite representação matricial.

Forma Quadrática: Q: $R^n \to R$

$Q(x_1, ..., x_n) = \sum a_{ij}x_ix_j$.

Como visto anteriomente: Q(x) = x^TA x , para uma matriz A simétrica.

Exemplo: Q: $R^2 \to R$: $Q(x_1, x_2) = a_{11}x_1^2 + a_{12}x_1x_2 + a_{22}x_2^2$

$$Q = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} a_{11} & \frac{1}{2}a_{12} \\ \frac{1}{2}a_{12} & a_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

Definição de formas quadráticas

Não existe termo independente. Logo Q(x = 0) = 0. x = 0 é um ponto de máximo, mínimo, ou nenhum dos dois?

Uma dimensão: Q(x) = ax^2

Se a > 0, x = 0 é um ponto de mínimo: y$\geq$0. Q positivo definido (P.D.).

Se a < 0, x = 0 é um ponto de máximo: y$\leq$0. Q negativo definido (N.D.).

Duas dimensões: y = Q(x_1, x_2) = x_1^2+ x_2^2

(x_1, x_2) $\neq$(0,0) $\Longrightarrow$y > 0: logo, Q é P.D.

y = - x_1^2 - x_2^2 : análogo: Q é N.D.

Se uma função quadrática pode assumir valores positivos e negativos, é dita indefinida.

Exemplo: y = x_1^2 - x_2^2

Se nunca é negativa, mas pode ser igual a 0 para x $\neq$0, é dita positiva semi-definida

Exemplo: y = x_1^2+ 2 x_1x_2 + x_2^2 = $(x_1+ x_2)^2$= 0 no ponto (1, -1).

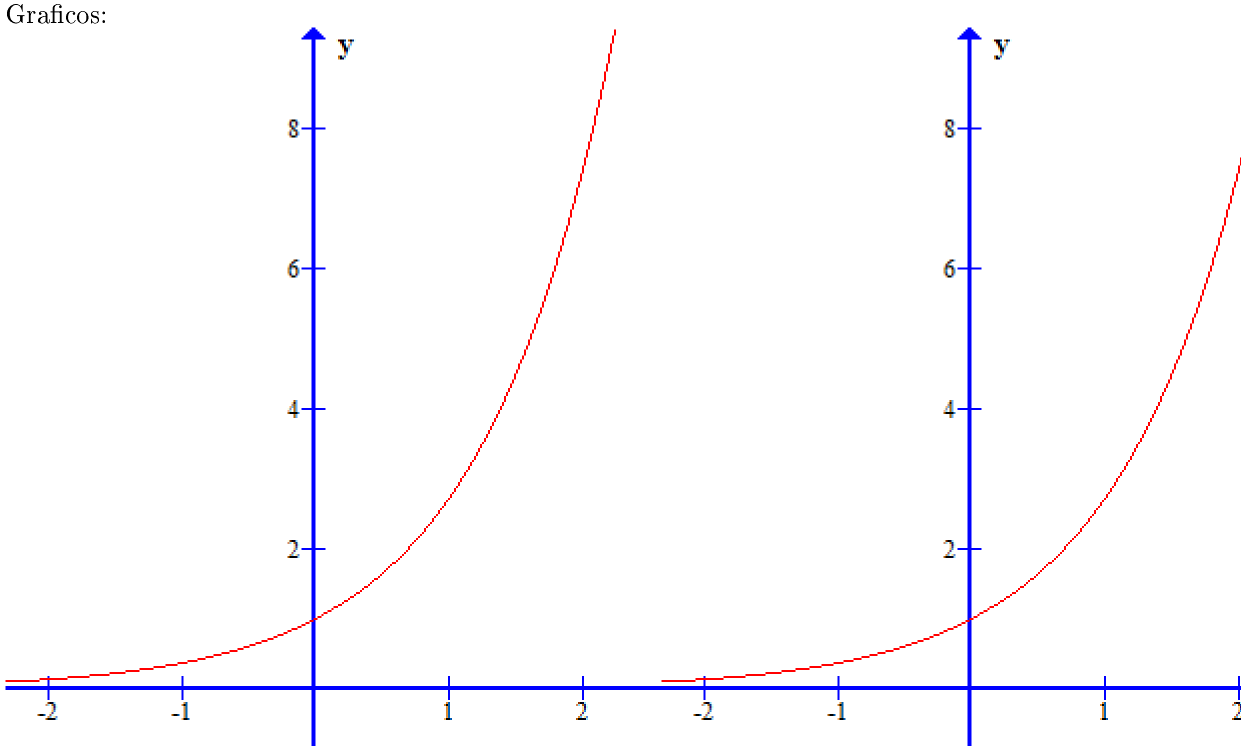
Análogo para negativa semi-definida.
Exemplo: -(x_1+ x_2)^2

Podemos definir a matriz A a partir das propriedades da forma quadrática $Q(x) = x^T Ax$, x$\in$R^m. Seja A simétrica. Então A é:

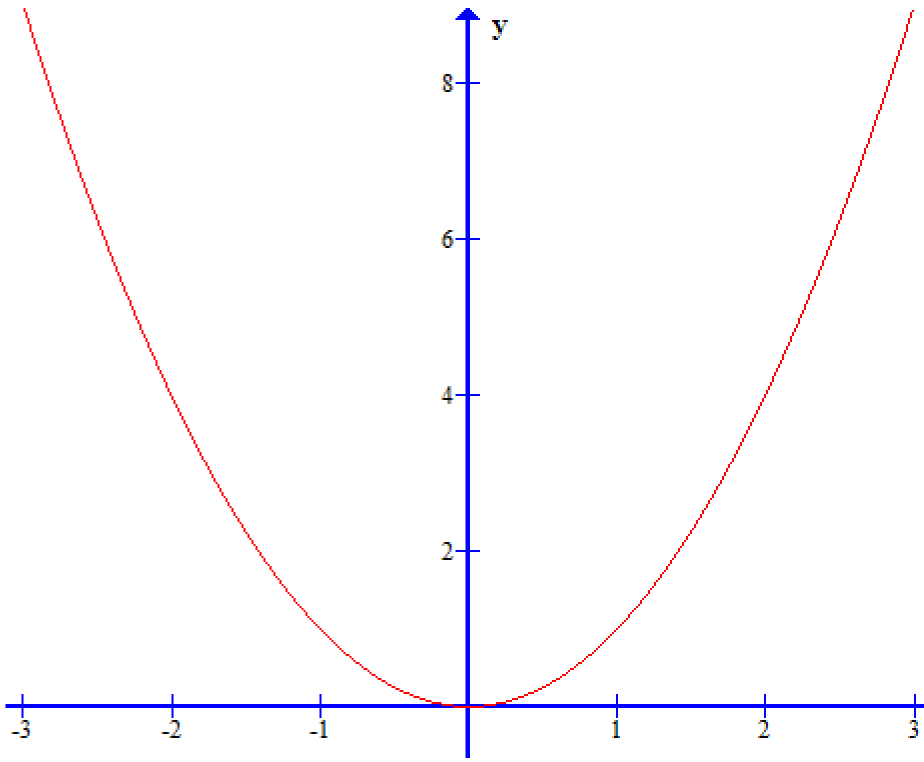
a) P.D se x^T A x $>$ 0 $\forall x \neq 0$ $\in$R^m.
b) P.S.D se x^T A x $\geq$ 0 $\forall x \neq 0$ $\in$R^m.
c) N.D se x^T A x $<$ 0 $\forall x \neq 0$ $\in$R^m.
d) N.S.D se x^T A x $\leq$ 0 $\forall x \neq 0$ $\in$R^m.
e) Indefinida caso não se encaixe em nenhum dos casos anteriores.

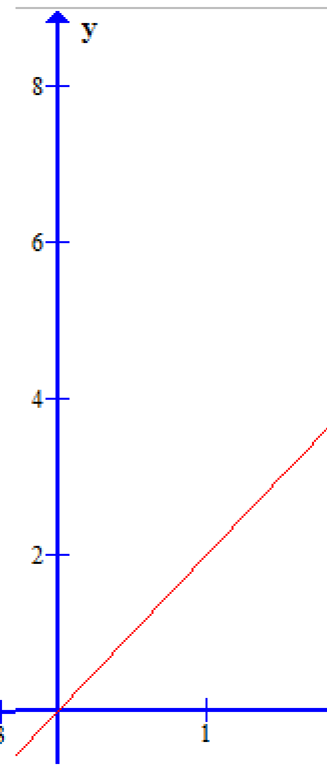
<u>Convexidade e condições de 2ª ordem.</u>

Graficos:



$y = e^x$ $\qquad$ $y' = e^x$





$y = x^2$ $\qquad$ $y' = 2x$

Objetivo: generalizar a derivada de 2ª ordem para dimensões superiores para avaliar convexidade/concavidade. Para tanto, vamos avaliar a definição da matriz Hessiana.

	Uma dimensão	*Váriias Dimensões*
Função convexa	$f'' > 0$	*H PD*
Função côncava	$f'' < 0$	*H ND*

Teste para Definição de Matriz

Menores: São os determinantes de submatrizes de uma matriz A qualquer, retirando determinadas linhas e colunas. Se forem retiradas linhas e colunas de mesmo índice (por exemplo, linha 1 e coluna 1, ou linhas 2 e 3 e colunas 2 e 3), teremos um menor principal líder.

Exemplo: A $= \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$

Menor principal de ordem 2: Det(A)

Menores principais de ordem 1: a_{11} e a_{22}

Exemplo 2: A $= \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$

Menor principal de ordem 3: Det (A)

Menores principais de ordem 2: $\det\left(\begin{bmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{bmatrix}\right)$, $\det\left(\begin{bmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{bmatrix}\right)$, $\det\left(\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}\right)$

Menores principais de ordem 1: a_{33} , a_{22} e a_{11}

Os menores principais mais usados , para cada ordem, são formados por eliminação das últimas linhas e colunas. São conhecidos como Menores Principais Líderes (MPL's):

a_{11}, $\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$, $\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$

Teorema: Seja $A_{(mxn)}$. Então:

1 - A é P.D se todos os MPL's são estritamente positivos

2 - A é N.D se os MPL's altenarem de sinal, começando com o sinal negativo: $A_1 < 0$, $A_2 > 0$, ...

Exemplos:

1 - $\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$ $A_1 = 2 > 0$, $A_2 =$ 2*2 - 1*1 $= 3 > 0$: P.D

2 - $\begin{bmatrix} -1 & -2 \\ 2 & 1 \end{bmatrix}$ $A_1 =$ -1 < 0, $A_2 =$ (-1)*1 - 2*(-2) $= 3 > 0$: N.D

3 - $\begin{bmatrix} 2 & 3 \\ 3 & 7 \end{bmatrix}$ $A_1 = 2 > 0$, $A_2 =$ 2*7 - 3*3 $= 5 > 0$: P.D

4 - $Q = x_1^2 + x_2^2 = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \underset{\substack{\downarrow \\ A}}{\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$

$A_1 = 1 > 0$, $A_2 = 1 \cdot 1 - 0 \cdot 0 = 1 > 0$: P.D

5 - $Q = - x_1^2 - x_2^2 = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \underset{\substack{\downarrow \\ A}}{\begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -x_1 & -x_2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$

$A_1 = -1 < 0$, $A_2 = (-1)(-1) - 0 \cdot 0 = 1 > 0$: N.D

6 - $Q = f(x_1 , x_1)$: Função de produção : côncava ?

$$H = \begin{bmatrix} \frac{d^2 f}{dx_1^2} & \frac{d^2 f}{dx_1 dx_2} \\ \frac{d^2 f}{dx_1 dx_2} & \frac{d^2 f}{dx_2^2} \end{bmatrix}$$

f côncava $\Longrightarrow$H é N.D $\Longrightarrow \frac{d^2 f}{dx_1^2} < 0$, $\frac{d^2 f}{dx_1^2} * \frac{d^2 f}{dx_2^2} - (\frac{d^2 f}{dx_1 dx_2})^2 > 0$

Para checar se a matriz é semi-definida , são necessários todos os menores principais

Teorema: $A_{(mxn)}$. A é P.S.D se e somente se todo M.P ≥ 0. A é N.S.D se e somente se todo M.P altena de sinal, começando com sinal negativo.

Matrizes Diagonais:

$$\begin{bmatrix} a_1 & 0 & \cdots & 0 \\ 0 & a_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & a_n \end{bmatrix}$$ Representa $a_1x_1^2 + \cdots + a_nx_n^2$

P.D se $a_i > 0\ \forall$i
N.D se $a_i < 0\ \forall$i
P.S.D se $a_i \geq 0\ \forall$i
N.S.D se $a_i \leq 0\ \forall$i

Otimização sem restrição

Seja y = F(x), x $\in R^n$. Se x^* é máximo local interior de F, então as derivadas parciais de F são igual a zero no ponto x^*.

Essa é uma condição necessária e não suficiente. Se a Hessiana no ponto x* for N.D, então é suficiente.

Condição de 1ª ordem

Ponto crítico:
R^1:$f'(x^*) = 0$
R^n: $\frac{df}{dx_i}(x_1^*, \cdots, x_n^*) = 0\ \forall$i
Ou seja, ∇f (x*) = 0.

Exemplo: Encontre os pontos críticos de $y = x^3 - y^3 + 9xy$
$\frac{dj}{dx} = 0 \Rightarrow 3x^2 + 9y = 0$
$\frac{dj}{dy} = 0 \Rightarrow -3y^2 + 9x = 0$
Soluções: (0, 0) e (3, -3) são pontos críticos

Condições de 2ª ordem

Precisamos avaliar a Hessiana em cada ponto crítico.
R^1: f"(x*) < 0: x^* é ponto de máximo
R^n: H é N.D: x^* é ponto de máximo

Se x^* é ponto critício e $D^2f(x^*)$ é N.D, então x^* é ponto de máximo local.
Se x^* é ponto critício e $D^2f(x^*)$ é P.D, então x^* é ponto de mínimo local.
Se x^* é ponto critício e $D^2f(x^*)$ é Indefinida(Nem N.D e nem P.D), então x^* é ponto de sela.

Essas são condições suficientes, mas não necessárias. As condições necessárias são mais fracas: H P.S.D (Min) ou N.S.D (Max):
Se x^* for ponto de máximo, então H(x^*) é N.S.D e ∇f (x^*) = 0.
Se $H(x^*)$ for ND e ∇f (x^*) = 0, então x^* é ponto de máximo

Exemplo: $y = x^3 - y^3 + 9xy$
pontos críticos : (0,0) e (3,-3)

$$H(x, y) = \begin{bmatrix} 6x & 9 \\ 9 & -6y \end{bmatrix}$$

H(0,0) = $\begin{bmatrix} 0 & 9 \\ 9 & 0 \end{bmatrix}$: A_1 = 0 , A_2= -81 < 0 : Indefinida. (0,0) é ponto de sela

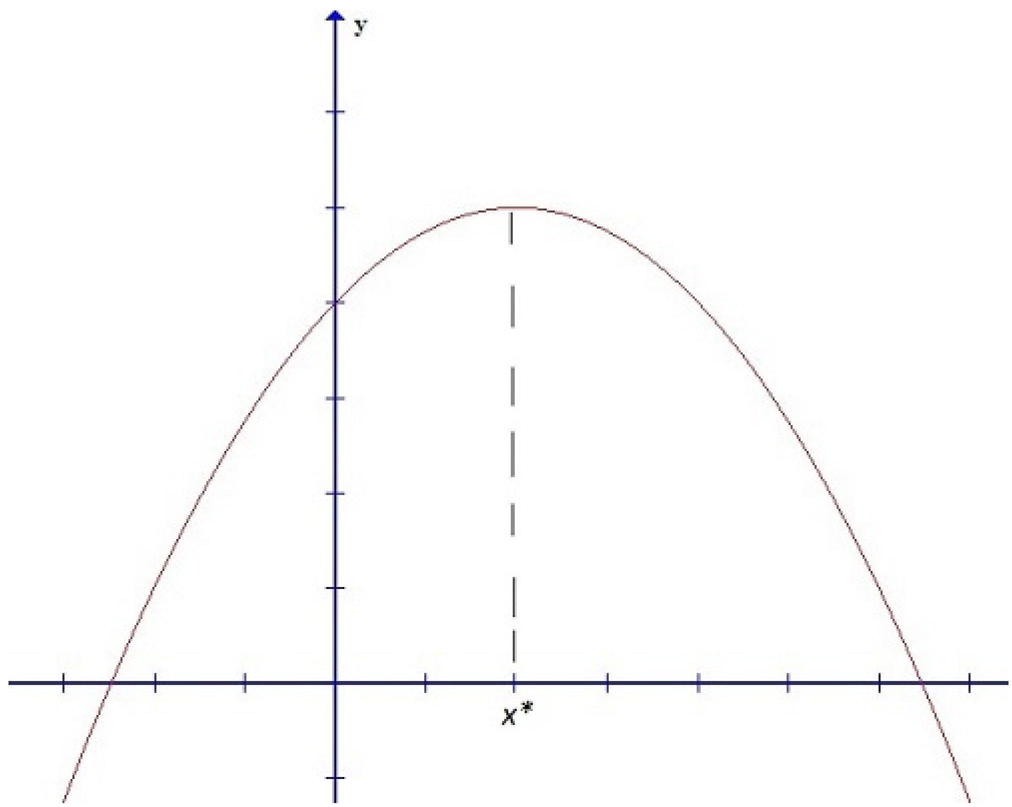
H(3,-3) = $\begin{bmatrix} 18 & 9 \\ 9 & 18 \end{bmatrix}$: A_1 = 18 > 0 , A_2= 18* 18 - 9*9 = 324 - 81 = 243 > 0 : P.D. (3,-3) é ponto de mínimo.

Máximo e Mínimo Global

Até aqui, consideramos máximos e mínimos locais.

Se f é côncava ∀x e $\exists x^*$ / $\nabla f(x^*) = 0$, então x^*é máximo global. D^2f $\leq$0.

Intuição: Derivada decrescente para todo x (função côncava tem derivada decrescente). Logo, não pode haver mais do que um ponto crítico (se a função é estritamente decrescente, ela pode passar pelo zero apenas uma vez). E como a função é côncava, o ponto crítico será de máximo local - único, e portanto global.



Exemplo: Maximização do Lucro.
f: função de produção
x_i: insumo i
w_i: preço do insumo i

p: preço do produto vendido

$$\max_{x_1, ..., x_n} \quad pf(x_1,..., x_n) - w_1x_1 - \cdots - w_nx_n$$

C.P.O: $p\frac{df}{dx_i}$ - $w_i = 0\ \forall i$

C.S.O: $P\begin{bmatrix} \frac{d^2f}{dx_I^2} & \cdots & \frac{d^2f}{dx_idx_n} \\ \vdots & \ddots & \vdots \\ \frac{d^2f}{dx_ndx_i} & \cdots & \frac{d^2f}{dx_n^2} \end{bmatrix} \leq 0$ (N.S.D)

Otimização com restrição

Problema Geral:

Max $f(x_1,..., x_n)$
sujeito às restrições $g_i(x_1,..., x_n) \leq b_i$, $i = 1, ..., k$
$h_j(x_1,..., x_n) = c_j$, $j = k + 1, ..., m$

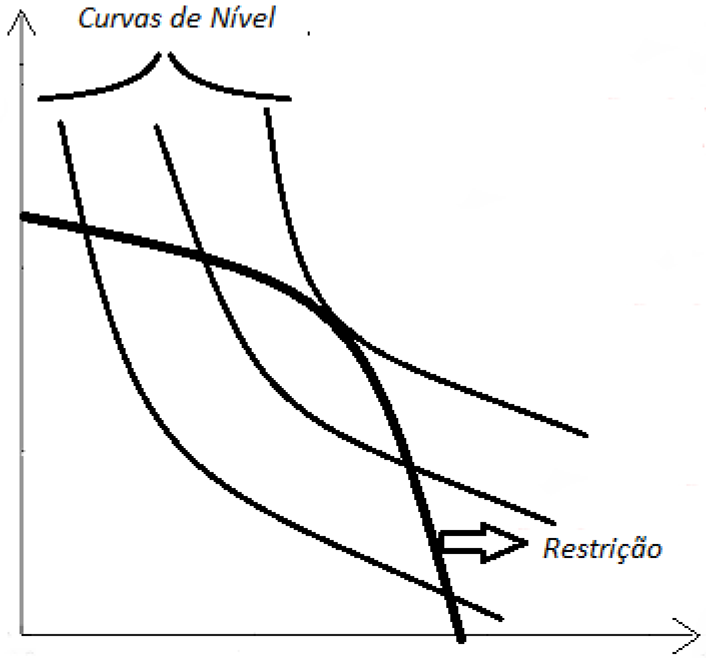
Ou seja, no total há m restrições: k restrições de desigualdade, $m - k$ restrições de desigualdade.

f: função objetivo (que se pretende maximizar ou minimizar)
g's: restrições de desigualdade
h's: restrições de igualdade

Exemplo: max $U(x_1,..., x_n)$
s.a: $p_1x_1 + ... + p_nx_n \leq M$
$x_i \geq 0\ \forall i$

Maximização com restrições de igualdade

Max $f(x_1, x_2)$
s.a $p_1x_1 + p_2x_2 = I$
(Na notação anterior, $h(x_1, x_2) = p_1x_1 + p_2x_2$ e $c = I$)



Curva de Nível mais alta não pode "cortar" a restrição, ou seria possível aumentar o valor de f. Logo, deve valer uma condição de tangência - ou seja, as inclinações das curvas de nível da restrição e da função objetivo devem ser iguais na solução do problema.

Como calcular a inclinação de uma curva de nível para uma função qualquer $a(x_1, x_2)$ associada a um valor b qualquer?
Curva de nível: $a(x_1, x_2) = b \ \forall x_1, x_2$
Derivada total:
$\frac{\partial a(x_1,x_2)}{\partial x_1} dx_1 + \frac{\partial a(x_1,x_2)}{\partial x_2} dx_2 = db = 0$

Ou seja: a variação total do lado esquerdo deve ser igual à variação total do lado direito, e essa última é igual a zero, pois é a variação de uma constante (ou seja, de um termo que não varia). Reorganizando:

$\frac{dx_2}{dx_1} = -\frac{\frac{\partial a(x_1,x_2)}{\partial x_1}}{\frac{\partial a(x_1,x_2)}{\partial x_2}}$

Ou seja: a inclinação dx_2/dx_1 da curva de nível é a razão das derivadas parciais.

Inclinação da curva de nível da função objetivo: $\frac{-\frac{df}{dx_1}}{\frac{df}{dx_2}}$

Inclinação da curva de nível da restrição: $\frac{-\frac{dh}{dx_1}}{\frac{dh}{dx_2}}$

Logo, a condição de tangência (inclinações iguais) encontrada acima pode ser escrita como : $\frac{\frac{df}{dx_1}}{\frac{df}{dx_2}} = \frac{\frac{dh}{dx_1}}{\frac{dh}{dx_2}}$

Reescreva como: $\frac{\frac{df}{dx1}}{\frac{dh}{dx1}} = \frac{\frac{df}{dx2}}{\frac{dh}{dx2}}$

Defina: μ= $\frac{\frac{df}{dx1}}{\frac{dh}{dx1}} = \frac{\frac{df}{dx2}}{\frac{dh}{dx2}}$

Então:

$\frac{\frac{df}{dx1}}{\frac{dh}{dx1}} = \mu \Rightarrow \frac{df}{dx1}$ - $\mu\frac{dh}{dx1} = 0$

$\frac{\frac{df}{dx2}}{\frac{dh}{dx2}} = \mu \Rightarrow \frac{df}{dx2}$ - $\mu\frac{dh}{dx2} = 0$

Acrescentamos mais uma variável (μ) ao problema, assim, ficamos com três variáveis e duas equações. Para resolver esse sistema, usamos a restrição como 3ª equação do sistema:

h(x_1, x_2) - c = 0

Outra forma de se chegar ao mesmo sistema é montar a função lagrangeano e derivá-la (igualando a zero, C.P.O) em relação a x_1, x_2 eμ:

L (x_1, x_2,\mu) = f(x_1, x_2) -μ(h(x_1, x_2) - c)

Derivada em relação a x_1: $\frac{df}{dx1}$ - $\mu\frac{dh}{dx1} = 0$

Derivada em relação a x_2: $\frac{df}{dx2}$ - $\mu\frac{dh}{dx2} = 0$

Derivada em relação aμ: h(x_1, x_2) - c = 0

As duas primeiras derivadas, em relação às variáveis originais do problema, podem ser escritas como:

$$\frac{\partial f(x_1,x_2)}{\partial x_1} = \mu\frac{\partial f(x_1,x_2)}{\partial x_1}$$
$$\frac{\partial f(x_1,x_2)}{\partial x_1} = \mu\frac{\partial f(x_1,x_2)}{\partial x_1}$$

Ou seja:

$$\nabla f\left(x^*\right) = \mu\nabla h\left(x^*\right)$$

Portanto, os gradientes da função objetivo f e da restrição h são paralelos na solução: um é múltiplo do outro.

As CPOs do Lagrangeano geram portanto todas as condições para a solução do problema de otimização sujeito a restrição de igualdade - ou seja, tanto a condição de tangência quanto a própria restrição.

Maximização com restrições de desigualdade

Na prática, o problema já foi resolvido: como veremos, ou a solução se comporta como em um problema sem restrição, ou se comporta como em um problema com restrição de igualdade. Trata-se apenas de organizar isso. Considere a versão mais simples do problema, com duas variáveis e apenas uma restrição, agora de desigualdade:

$\underset{x_1,x_2}{Max}\, f\left(x_1, x_2\right)$

sujeito à restrição $g\left(x_1, x_2\right) \leq c$

Montamos o lagrangeano da mesma forma, usando o multiplicador de lagrange λ:

$$L(x_1, x_2, \lambda) = f(x_1, x_2) - \lambda [g(x_1, x_2) - c]$$

A solução pode estar em duas regiões:

I) Solução com restrição ativa: a restrição vale com igualdade: $g(x_1, x_2) = c$
II) Solução com restriição inativa: a restrição vale com desigualdade estrita: $g(x_1, x_2) < c$

No primeiro caso, devem valer as mesmas condições já vistas para solução do problema da maximização com restrição de igualdade:

$$I) \quad \begin{cases} \nabla f(x_1, x_2) = \lambda \nabla g(x_1, x_2) \\ g(x_1, x_2) = c \end{cases}$$

Além disso, há apenas um detalhe adicional: o multiplicador não pode assumir um sinal qualquer. A ideia é a seguinte: os gradientes de f e g devem apontar para no mesmo sentido (não basta ter a mesma direção, ou seja, não basta ser paralelos). Caso contrário, se apontassem em sentidos opostos, seria possível aumentar o valor da função f movendo a solução para dentro do conjunto admissível, e portanto não teríamos uma solução. Como implicação, o multiplicador deve ser não-negativo: $\lambda \geq 0$ (note que o sinal da desigualdade na restrição não pode mudar, ou seria necessário mudar também o sinal do multiplicador!).

No segundo caso, a restrição não restringe de fato a escolha do ponto de máximo de f: seria possível aumentar ou diminuir um pouco x_1 e/ou x_2 e ainda assim respeitar a restrição $g(x_1, x_2) < 0$, mas isso não é feito, o que significa que a solução não é afetada pela presença da restrição. Logo, estamos resolvendo um problema de maximização sem restrição, o que equivale a fazer $\lambda = 0$ no lagrangeano para obter as mesmas condições já vistas para maximização sem restrição:

$$II) \quad \nabla f(x_1, x_2) = 0$$

Como escrever condições para o ótimo que permitam esses dois casos (restrição ativa ou inativa)? Basta notar que, em qualquer caso, deve valer a condição $\lambda \cdot [g(x_1, x_2) - c] = 0$: quando a restrição é ativa, $[g(x_1, x_2) - c] = 0$; quando é inativa, $\lambda = 0$. Lembre-se ainda de que devemos impor $\lambda \geq 0$.

As condições para o ótimo podem ser escritas como:

$$\begin{cases} \nabla f(x_1, x_2) = \lambda \nabla g(x_1, x_2) \\ g(x_1, x_2) \leq c \\ \lambda \cdot [g(x_1, x_2) - c] = 0 \\ \lambda \geq 0 \end{cases}$$

Exemplo: $\underset{x_1, x_2}{Max}\ f(x_1, x_2)$
sujeito a $x_1 \leq 0$
Lagrangeano:

$L(x_1, x_2, \lambda) = f(x_1, x_2) - \lambda x_1$
Condições de primeira ordem:
$\left(\frac{\partial f(x_1,x_2)}{x_1}, \frac{\partial f(x_1,x_2)}{x_2}\right) = \lambda \cdot (1, 0)$
$x_1 \leq 0$
$\lambda x_1 = 0$
$\lambda \geq 0$

Note que, em uma solução interior ($x_1 > 0$), devemos necessariamente ter $\lambda = 0$, e portanto $\frac{\partial f(x_1,x_2)}{x_1} = 0$. Em uma solução de canto, porém, $\frac{\partial f(x_1,x_2)}{x_1} = \lambda > 0$.

Função Valor e Teorema do Envelope

A função valor é simplesmente o valor que a função assume na solução. Por exemplo: é a utilidade máxima atingida no problema do consumidor.
$\underset{x_1,x_2}{Max}\ f(x_1, x_2)$
sujeito a $h(x_1, x_2) \leq c$
Encontramos a solução $[x_1^*(c), x_2^*(c)]$. Note que a solução depende dos parâmetros do problema; isso estava sempre implícito, e agora a notação mostra explicitamente. Por exemplo: na solução do problema do consumidor, a demanda depende dos parâmetros do problema: preços e renda: $x^*(p_1, p_2, M)$.
Montamos então a função valor v, que dependerá dos parâmetros:

$$v(c) = f(x_1^*(c), x_2^*(c))$$

Por exemplo: para o consumidor com utilidade Cobb-Douglas $u(x_1, x_2) = x_1x_2$, as demandas são $x_1 = \frac{M}{2p_1}$ e $x_2 = \frac{M}{2p_2}$. Logo, a função valor é:

$$v(p_1, p_2, M) = x_1^*x_2^* = \frac{M}{2p_1}\frac{M}{2p_2} = \frac{M^2}{4p_1p_2}$$

Note ainda que, como $\lambda \cdot [h(x_1, x_2) - c] = 0$ (mesmo que seja uma restrição de desigualdade) , a função valor é escrita a partir do lagrangeano - note que o multiplicador também depende do parâmetro na solução:

$$v(c) = f(x_1^*(c), x_2^*(c)) - \lambda^*(c) \cdot [h(x_1^*(c), x_2^*(c)) - c]$$

Podemos então perguntar como a função valor responde a uma mudança nos parâmetros. Ou seja, podemos calcular a derivada $v'(c)$. A princípio, essa derivada pode parecer longa, pois o parâmetro c aparece como argumento de várias funções dentro da função valor.

É possível simplificar essa derivada: pelo teorema do envelope, só precisamos olhar o impacto direto de c sobre o valor da função, ignorando o impacto indireto, ou seja, ignorando o impacto de c sobre os argumentos x_1, x_2, λ:

$$v'(c) = \lambda$$

Isso ocorre porque as demais derivadas somam zero: elas apenas reproduzem a derivada primeira, que deve ser igual a zero na solução. Se calcularmos diretamente a derivada, vamos encontrar:

$$v'(c) = f_1 \cdot \frac{dx_1^*}{dc} + f_2 \cdot \frac{dx_2^*}{dc} - \frac{d\lambda^*}{dc}\left[h\left(x_1^*\left(c\right), x_2^*\left(c\right)\right) - c\right] - \lambda^*\left(c\right)\left[h_1 \cdot \frac{dx_1^*}{dc} + h_2 \cdot \frac{dx_2^*}{dc} 1\right] + \lambda^*\left(c\right)$$

À exceção do último termo, todos os demais somam zero devido às condições de primeira ordem.

Interpretações:

1- no problema do consumidor, o multiplicador é a utilidade marginal da renda.

2- no problema de minimização de custo da firma, o multiplicador é o custo marginal.