# PROGRAMAÇÃO NÃO LINEAR - P.N.L.

*Fernando Pereira Tostes*
*Mestrado (Wharton School of Finance and Commerce)*
*Cursando Doutorado em Ciências Contábeis pela FEA USP*

## I. INTRODUÇÃO

Programação linear surgiu em meados deste século e representou um avanço em técnicas de otimização, já que as restrições podem entrar num problema na forma de desigualdades, entre as quais as mais comuns são as de não-negatividade. Entretanto, uma premissa-chave em programação linear é que *todas* as funções envolvidas (a função objetivo e as restrições) *são lineares*.

Embora essa premissa seja adequada para grande número de problemas práticos, isso nem sempre acontece. Muitos economistas acreditam até, que um certo grau de *não-linearidade* é a regra, ao em vez da excessão.

Chiang apresenta o problema de programação não-linear na forma de maximização assumindo o formato geral conforme abaixo:

**TABELA 1**

**PROBLEMA NÃO-LINEAR**

Maximizar $\pi = f\ (x_1, x_2, \ldots\ldots x_n)$

sujeito a: $g^1(x_1, x_2, \ldots\ldots x_n) \leq r_1$

$g^2(x_1, x_2, \ldots\ldots x_n) < r_2$

.....................................

$g^m(x_1, x_2, \ldots\ldots x_n) \leq r_m$

e $x_j > 0\ \ (j = 1,2,\ldots\ldots n)$

onde $f$ é uma função côncava definida em $R^n$ a valores reais, e $g^1, g^2, \ldots g^n$, são funções convexas, também definidas em $R^n$ a valores reais.

Numa forma mais compacta, o problema de programação não-linear também pode ser descrito como achar

$\mathbf{x} = (x_1, x_2, \ldots\ldots x_n)$ de forma a maximizar $f(x)$,

sujeito a $g(x) <= b$, para $i = 1,2,\ldots\ldots m$

e $x => 0$,

onde $f(x)$, e $g(x)$ são funções das $n$ variáveis de decisão,

$g(x)= (g^1(x), \ldots, g^m(x))$ e

$b= (r_1, r_2, \ldots., r_m)$.

Tal como em programação linear, o problema acima possue três componentes:

- uma função objetivo
- um conjunto de $m$ restrições, e
- um conjunto de condições de não-negatividade

para as $n$ variáveis em questão. Igualmente se supõe que todas as funções sejam diferenciáveis, e $m$ pode ser maior, igual ou menor que $n$.

Nenhum algoritimo irá solucionar cada problema especifico usando esse formato genérico. Mesmo assim, tem havido progressos em vários casos especiais do problema, onde se estuda e estima o comportamento das respectivas funções.

Em problemas econômico-financeiros, por exemplo, existem várias razões pela qual uma função linear se torna não-linear. É comum supor-se em problemas de programação linear que o lucro unitário de um produto seja constante. Isso nem sempre ocorre, podendo ser uma função decrescente do volume de produção, isto é, quanto maior o volume, menor o lucro unitário. Assim a função objetiva linear $f(\mathbf{x}) = (c_1 x_1 + C_2 x_2 + \ldots\ldots c_n x_n)$ seria trocada por outra não-linear do tipo $P(\mathbf{x}) = (c_1(x_1)x_1 + C_2(x_2)x_2 + \ldots\ldots c_n(x_n)x_n)$ onde $c_j(x_j)$ traduz uma função crescente da variável $x_j$. Da mesma forma poderiam haver alterações no âmbito das restrições. Por exemplo, o insumo $i$ (energia elétrica) decresce à medida que aumenta o nível de produção de $f$, porque as últimas unidades são processadas mais rapidamente que as anteriores, isto é, usando menos tempo de máquina por unidade. Em programação linear, o coeficiente $c_{ij}$ seria constante, mas no caso acima tal premissa ficaria prejudicada. O mesmo ocorreria se o coeficiente $c_{ij}$ depender do insumo $i$ *mais* o insumo $k$, incluindo um termo adicional nas restrições, que por sua vez englobam as variáveis $x_j$ e $x_k$, eliminando novamente a linearidade.

Em tais casos, a formulação do problema de modo não-linear é recomendada, e com isso várias conveniências aplicáveis à solução linear não podem ser usadas. Entre elas o conforto de poder contar com programas de computador de fácil manuseio, permitindo soluções rápidas em diversas alternativas.

## II. EMBASAMENTO TEÓRICO DE P.N.L.

Dado o problema anterior

**TABELA 1**

**<u>PROBLEMA NÃO-LINEAR (PNL)</u>**

$$
\begin{aligned}
\text{Maximizar } \pi = f\ & (x_1, x_2, \ldots\ldots x_n) \\
\text{sujeito a }\ & g^1(x_1, x_2, \ldots\ldots x_n) \le r_1 \\
& g^2(x_1, x_2, \ldots\ldots x_n) \le r_2 \\
& \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\
& g^m(x_1, x_2, \ldots\ldots x_n) \le r_m \\
\text{e }\ & x_j \ge 0 \ \ (j = 1,2,\ldots\ldots n)
\end{aligned}
$$

onde $f$ é uma função côncava definida em $R^n$ a valores reais, e $g^1$, $g^2$,... $g^n$, são funções convexas, também definidas em $R^n$ a valores reais.

Vamos supor que tal problema seja *fortemente consistente*, isto é, existe um ponto $x_0$, pertencente a $R^n$, tal que $g^i(x_0) < 0$, para i = 1,2,...., m, e $x_0 > 0$. Uma condição desse tipo é chamada *condição de Slater*. Este quesito sobre o problema original assegura a *qualificação das restrições* que é uma hipótese básica para a validade do teorema de KARUSH-KUHN-TUCKER, que enunciaremos a seguir.

Lembramos que para uma função $f$ definida em $R^n$ a valores reais e diferenciável, denomina-se *gradiente de $f$* ao vetor cujas coordenadas são as derivadas parciais da função $f$, isto é,

$$\nabla f(x) = (\partial_1 f(x),\ \partial_2 f(x),\ \ldots\ \partial_n f(x)),$$

onde

$\partial_j f(x)$ é a derivada parcial de $f$ com relação à variável $j$, sendo $j = 1, 2, \ldots, n$.

***Teorema (KARUSH-KUHN-TUCKER):***

Suponha que as funções $f$ e $g^1, \ldots, g^m$ são a valores reais, definidas em todo $R^n$ e diferenciáveis. $f$ é côncava e as $g^1, \ldots, g^m$ são convexas. Logo um vetor $x^*$ pertencente a $R^n$ é um ótimo do "PROBLEMA PNL" se, e somente se, existem $\lambda^*$ pertencente a $R^m$ e $\mu^*$ pertencente a $R^n$, tais que

i ) $\nabla f(x^*) - \Sigma\, \lambda_i^* \nabla g^i(x^*) = (\mu_1^*, \mu_2^*, \ldots, \mu_n^*)$

ii) $\lambda_i^* g^i(x^*) = 0$ para $i = 1, 2, \ldots, m$
$\mu_j^*\, x_j^* = 0$ para $j = 1, 2, \ldots, n$

iii) $\lambda^* \geq 0$
$\mu^* \geq 0$

A condição (i) é chamada de *condição lagrangiana* e é a que traz mais informações sobre o problema.

A condição (ii) é denominada *condição de complementariedade* e indica que se $x^*$ satisfaz alguma restrição com folga, isto é, se alguma $g^i(x^*) < 0$ então o multiplicador correspondente $\lambda_i^*$ deve ser nulo. Portanto, as variáveis não-nulas do multiplicador correspondem àquelas restrições tais que $g^i(x^*) = 0$.

A condição (iii) é chamada a *condição de viabilidade dual*, que significa que os multiplicadores $\lambda^*$ e $\mu^*$ devem ser não-negativos.

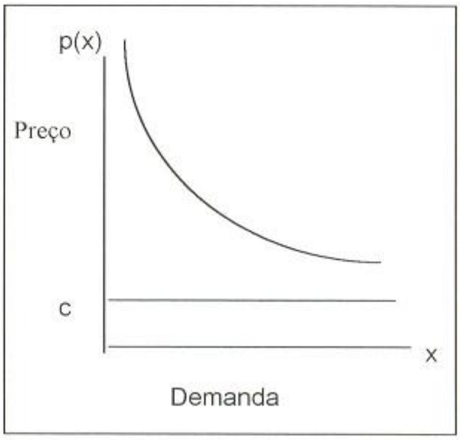
## III. APLICAÇÕES TÍPICAS DE P.N.L.

### *1. MIX DE PRODUÇÃO COM ELASTICIDADE DE PREÇO*

Um problema clássico de programação linear é o de encontrar o *mix ótimo* de produção, de modo a *maximizar* a margem de contribuição global da empresa, mas considerando as diversas limitações de recurso para para fabricar os artigos. Em geral há um lucro unitário associado a cada produto, permitindo que a função objetiva seja linear. No entanto, observa-se com freqüência nesse tipo de problema, fatores que introduzem *não-linearidades* na função objetiva, como *elasticidade de preço*. Nesse caso, a quantidade a ser vendida é uma função inversa do preço cobrado, e a curva de *preço-demanda* aparece como na FIGURA 1 abaixo, onde $p(x)$ é o preço desejado para vender $x$ unidades.

**FIGURA 1**

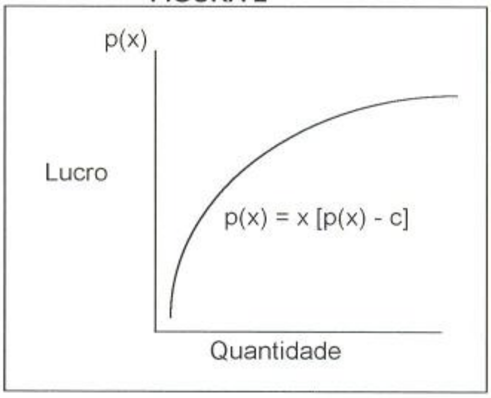
c = custo unitário




Se o custo unitário do produto é fixo em (c), a margem de contribuição para produzir e vender x unidades seria dada pela função não-linear abaixo, e ilustrada na FIGURA 2.

$$p(x) = x.p(x) - c(x)$$

**FIGURA 2**




**Função da Margem de Contribuição**

Cada um dos $n$ artigos da empresa terá uma função similar, $P_j(x_j)$ para produzir e vender $x_j$ unidades do produto $j$, sendo ($j$ = 1, 2, 3, ......, $n$). Logo, a função objetiva genérica será uma somatória de funções não-lineares.

$$\boxed{f(x) = \Sigma\, P_j(x_j)}$$

Existem outras razões, pela qual podem surgir não-linearidades. Por exemplo, o *custo marginal* unitário tende a cair com o aumento da quantidade produzida, em função do *efeito da curva de aprendizagem*, um incremento de eficiência produtiva à medida que aumenta a experiência. Mas pode também aumentar, pelo ônus adicional de encargos sociais, custo de manutenção, armazenagem, e outros inerentes ao aumento de produção.

Nas funções de restrição $g_i(x)$ também podem ocorrer situações semelhantes de não-linearidade, se houver digamos, uma restrição orçamentária ao custo de produção total, e o custo marginal mostrar um comportamento tal como descrito acima. $g_i(x)$ será não-linear sempre que o uso dos insumos correspondentes não forem exatamente proporcionais aos níveis de produção dos respectivos artigos.

## 2. SELEÇÃO DE CARTEIRA COM RISCO

Atualmente é comum a administração de grandes carteiras de ações e títulos de renda fixa através de computadores, que empregam programação não-linear para orientar as decisões de compra e venda. Como os investidores se baseiam conjuntamente com o *retorno (ganho) e o risco (oscilação de preço)* de seus patrimônios, programação não-linear é usado para organizar uma carteira que, dentro de certas premissas, produza uma uma combinação ótima entre esses dois fatores.

Imaginemos que um administrador deseje incluir *n* ações numa carteira, e que as variáveis de decisão $x_j$ (j= 1, 2, 3...... *n*) sejam o número de ações *j* a serem incorporadas. $\mu_j$ e $\sigma_{jj}$ são respectivamente a *média* e *variança (estimadas)* do retorno sobre cada ação *j*, onde $\sigma_{jj}$ mede o risco de cada ação. Para **i = 1,2,3.... *n***, sendo $i \neq j$, vamos denominar $\sigma_{ij}$ a *covariança* do retorno de uma ação i e outra *j*. O *valor esperado* R (x) e a variança V (x) do retorno global da carteira é obtido por

$$\boxed{\begin{aligned} \mathbf{R}\,(\mathbf{x}) &= \sum \mu_j\, x_j \\ \mathbf{V}\,(\mathbf{x}) &= \sum\sum \sigma_{ij}\, x_i\, x_j \end{aligned}}$$

onde **V (x)** mede o risco associado com a carteira. O truque para montar a combinação perfeita entre esses dois fatores é juntá-los na função objetiva a ser maximizada,

$$\boxed{f(x) = \mathbf{R}(x) - \beta \mathbf{V}(x)}$$

onde o parâmetro $\beta$ é uma constante não-negativa que reflete a combinação preferida do investidor entre risco e retorno. Assim, se $\beta=0$, o risco deve ser ignorado por completo, enquanto um alto valor de $\beta$ implica em dar um alto peso ao fator ao risco (ou seja, maximizar $-\mathbf{V}(\mathbf{x})$).

Portanto, o modelo completo de programação não-linear para o problema seria

$$\text{maximizar } f(x) = \sum \mu_j x_j - \beta \sum\sum \sigma_{ij}\, x_i\, x_j$$

sujeito a $\sum \mathbf{P}_j\, \mathbf{x}_j \leq \mathbf{B}$

e $x_j \geq 0$, para $j = 1, 2, 3, \ldots.. n$,

onde $\mathbf{P}_j$ é o preço de cada ação *j* e **B** é o total orçado para a carteira.

Conforme a *função-utilidade* do investidor (mede o valor relativo de diferentes retornos para o investidor) pode-se demonstrar que uma solução ótima para esse problema de programação não-linear *maximiza* a *utilidade esperada* (desejada) do investidor.[1]

Uma desvantagem da formulação acima é que, dado R(x) e V(x) serem relativamente impossíveis de medir, é um tanto difícil de escolher um valor apropriado para $\beta$. Por isso, ao em vez de se trabalhar com apenas um $\beta$, na prática usa-se um enfoque de programação

[1] Veja Bazaraa, Mokhtar S., e C.M.Shetty; "Nonlinear Programming: Theory and Algorithms"; Wiley, New York, 1979.

*paramétrica* (não-linear) para produzir a solução ótima como função de $\beta$, sobre uma gama de valores de $\beta$. O passo seguinte é o exame de R(x) e V(x) para estas soluções ótimas, em determinado $\beta$, e então escolher a solução que parece fornecer a melhor combinação entre essas duas quantidades. Esse procedimento é chamado de *produzir soluções na fronteira eficiente* de um gráfico bidimensional de R(x), V(x) pontos possíveis *x*. A razão é que o ponto R(x), V(x) ótimo (para um dado $\beta$), fica na fronteira do conjunto de pontos viáveis. E mais ainda, cada *x* ótimo é *eficiente* no sentido que nenhuma outra solução viável será tão boa com uma medida (R ou V) e melhor com outra medida (menor V ou maior R).

## IV. CASO PRÁTICO

Vamos agora aplicar a teoria da seção II ao caso particular da seção III.2, tomando como exemplo os dados da TABELA 2 abaixo.

**TABELA 2**

| AÇÃO | $\mu_j$ | $P_j$ \$ |
|---|---|---|
| Acesita (ON) | 3 | 5,00 |
| Brahma (PN) | 7 | 4,00 |
| B.Brasil (PN) | 10 | 2,00 |

onde

$\mu_j$ é a média estimada da rentabilidade de cada uma das ações. $\mu_j = 3$, por exemplo significa uma rentabilidade estimada para o período de 200%. $P_j$ é o preço de compra estimado da ação.

A medida de risco é obtida pela variança ($\sigma_{ij}$) de cada ação acima. E para cada uma computa-se a covariança ($\sigma_{ij}$) entre o retorno da ação *i* e *j*, montando-se a matriz abaixo.

$$\begin{matrix} \sigma_{11} & \sigma_{12} & \sigma_{13} \\ \sigma_{12} & \sigma_{22} & \sigma_{23} \\ \sigma_{13} & \sigma_{23} & \sigma_{33} \end{matrix}$$

onde

$\sigma_{11}$ é a variança do retorno da Acesita; $\sigma_{22}$, ídem para Brahma, e $\sigma_{33}$, ídem para Banco do Brasil (a covariança de uma variável com ela mesma é a sua própria variança).

$\sigma_{12}$ é a covariança entre o retorno da Acesita com Brahma.

$\sigma_{13}$, ídem para Acesita e Banco do Brasil

$\sigma_{21}$, ídem para Brahma e Acesita

$\sigma_{23}$, ídem para Brahma e Banco do Brasil

$\sigma_{31}$, ídem para Banco do Brasil e Acesita

$\sigma_{32}$, ídem para Banco do Brasil e Brahma

Para a matriz de covarianças foram arbitrados os números abaixo

$$\begin{matrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{matrix}$$

Dado um orçamento B = R\$500.000, o problema pode ser estruturado na forma abaixo:

$$\textbf{Maximizar } 3x_1 + 7x_2 + 10x_3 - \beta\,(x_1^2 + 2x_2^2 + x_3^2)$$

$$\text{sujeito a} \quad 5x_1 + 4x_2 + 2x_3 \leq 500.000$$

$$\text{e} \quad x_1, x_2, x_3 \geq 0$$

A solução ótima -- $x^* = (x_1^*, x_2^*, x_3^*)$ deverá satisfazer as condições (*i*), (*ii*) e (*iii*) do Teorema de Karush-Kuhn-Tucker. Processa-se os cálculos pertinentes, e toma-se $\alpha^* = (\alpha_1^*, \alpha_2^*, \alpha_3^*)$, o multiplicador correspondente às restrições de não-negatividade, e $\lambda^* \in R$, o multiplicador de restrição orçamentária. Temos então as equações das condições Karush-Kuhn-Tucker

(i) $3 - 2\beta x_1^* - 5\lambda^* = \alpha_1^*$ (1)

$7 - 4\beta x_2^* - 4\lambda^* = \alpha_2^*$ (2)

$10 - 2\beta x_3^* - 2\lambda^* = \alpha_3^*$ (3)

*(ii)* $\lambda^* (5x_1 + 4x_2 + 2x_3 - 500.000) = 0$ (4)

$\alpha_1^* x_1^* = 0$ (5)

$\alpha_2^* x_2^* = 0$ (6)

$\alpha_3^* x_3^* = 0$ (7)

*(iii)* $\lambda^* \geq 0$

$\alpha^* \geq 0$

De posse de todos os dados, passa-se ao cálculo da solução ótima como função da constante $\beta$ (grau de aversão ao risco do investidor). Presumindo que o gasto total não atinja o valor limite do orçamento -- R$500.000 --, a equação (4) implica que $\lambda^*$ (o multiplicador de restrição orçamentária) será igual a zero. Isso simplifica as equações (1), (2) e (3), transformando-as em

$3 - 2\beta x_1^* = \alpha_1^*$ (8)

$7 - 4\beta x_2^* = \alpha_2^*$ (9)

$10 - 2\beta x_3^* = \alpha_3^*$ (10)

Vamos agora multiplicar os dois membros da equação (8) por $x_1^*$, e obteremos

$3x_1^* - 2\beta (x_1^*)^2 = \alpha_1^* x_1^*$ (11)

A equação (5) implica que o membro da direita da expressão (11) é igual a zero. Logo,

$3x_1^* - 2\beta (x_1^*)^2 = 0$ (12)

Para facilitar o desenho da solução supomos que $x_1^* > 0$, o que significa que a quantidade de ações a ser comprada será sempre positiva. Tal suposição nos permite simplificar a variável $x_1^*$ da equação (12), obtendo então

$x_1^* = 3/2\beta$

Calculamos agora $x_2^*$ seguindo os mesmos passos executados para o cômputo de $x_1^*$.

Multiplica-se a equação (9) por $x_2^*$, obtendo-se assim

$7x_2^* - 4\beta (x_2^*)^2 = \alpha_2^* x_2^*$ (13)

Pelo mesmo raciocínio anterior, e usando a equação (6), temos que o membro direito de (13) é zero, obtendo-se

$7x_2^* - 4\beta (x_2^*)^2 = 0$ (14)

Novamente simplifica-se $x_2^*$, supondo que é não-nulo, para obter então

$x_2^* = 7/4\beta$

Repetindo o mesmo processo para $x_3^*$, obteremos $x_3^* = 5/\beta$.

## ANÁLISE DOS RESULTADOS

A solução ótima assim obtida é $x^* = 1/\beta\,(3/2; 7/4; 5)$ que é uma função do grau de aversão ao risco do investidor ($\beta$). Portanto, **quanto maior for o valor de $\beta$, menor será o investimento em cada ação e para valores muito pequenos de $\beta$ maior será a quantidade investida.**

O investidor irá comprar uma quantidade maior de ações do Banco do Brasil, porque conforme os dados das páginas 9 e 10, sua variança é igual a 1, a rentabilidade estimada é a mais alta (10, ou seja 900%) e o preço o mais baixo ($2). Conforme afirmamos na página 8, a expressão da solução em função de $\beta$ nos diz que a direção da solução ótima será sempre a mesma, isto é, para qualquer valor de $\beta$ o investidos sempre comprará uma maior quantidade de ações do Banco do Brasil. Posto de outra forma, a proporção entre todas as coordenadas do portfólio ótimo será a mesma.

## V. COMENTÁRIO FINAL

Em problemas de otimização é comum nos depararmos com um comportamento não-linear que deve ser levado em consideração. Em comparação ao método Simplex indicado para Programação Linear, não existe um algoritmo único que possa ser usado para resolver *todos* os problemas da PNL. No entanto, tem havido progresso em casos especiais como programação quadrática, programação convexa e não-convexa, incluindo o desenvolvimento *softwares* de uso geral para computador tipo main frame.

Para a contabilidade, o uso de PNL é limitado.

- a maior parte da *contabilidade financeira* está relacionada à mensuração e evidenciação de fenômenos econômicos, na qual não se emprega PNL.

- em *contabilidade gerencial*, a PNL tem utilidade nos problemas que objetivam apontar a alocação ótima de recursos. Este tipo de tarefa geralmente é designada a engenheiros, por requerer maior familiaridade com as ferramentas matemáticas necessárias.

A teoria de Programação Não-Linear é uma ferramenta interessante, mas que poucas pessoas tem um domínio suficiente para usá-la correta e habitualmente no trabalho. Mesmo entre os engenheiros, não são muitos os que são capazes de resolver problemas de PNL sem recordar nos livros. A base de sua maior utilidade está no campo de finanças, distribuição de recursos (problemas de transporte), mercadologia e administração da produção, casos nos quais a não-linearidade da função-objetivo se aproxima mais à realidade.