

Informações trade-off em problemas de otimização multiobjetivo

Jesus Ossian da C. Silva, Paulo Roberto Oliveira,

Programa de Engenharia de Sistemas e Computação-PESC, COPPE, UFRJ

21941-972, Rio de Janeiro, RJ

E-mail: jossian@cos.ufrj.br, poliveira@cos.ufrj.br,

Resumo: *Em problemas de otimização multiobjetivo temos o objetivo definido por k funções, onde as funções componentes do objetivo são conflitantes entre si. O conjunto solução do problema de otimização multiobjetivo é desprovido de relação de ordem entre seus elementos, o que dificulta a determinação de uma solução desejável. Uma alternativa para se determinar uma melhor solução é o uso de taxas trade-off as quais expressam a real quantidade de ganho e perda quando mudamos de uma solução para outra. Os multiplicadores de Karush-Kunh-Tucker podem ser interpretados como taxas trade-off. Nesse trabalho estudamos uma metodologia para tomada de decisões em problemas de otimização multiobjetivo de acordo com [1].*

Palavras-chave: *multiobjetivo, trade-off, Karush-Kunh-Tucker.*

1 Introdução

Definimos o problema o problema de otimização multiobjetivo.

Problema 1 (POM) *Minimizar* $(f_1(x), f_2(x), \dots, f_k(x))$,
Sujeito a $x \in S \subset \mathbb{R}^n$.

onde S é denominado conjunto viável do problema.

Consideremos o seguinte exemplo [3] de POM.

Exemplo 1 *minimizar* $((x_1 - 3)^2 + (x_2 - 2)^2, x_1 + x_2, x_1 + 2x_2)$
sujeito a $-x_1 \leq 0$
 $-x_2 \leq 0$.

Abaixo definimos o conjunto solução do POM.

Definição 1 (Pareto ótimo) x^* é solução Pareto ótimo do POM se não existe outro $x \in S$ tal que $f_i(x) \leq f_i(x^*)$ para todo $i = 1, \dots, k$ e $f_i(x) < f_i(x^*)$ para no mínimo um índice i .

Definição 2 (Pareto ótimo fraco) $x^* \in S$ é Pareto ótimo fraco se não existe outro $x \in S$ tal que $f_i(x) < f_i(x^*)$ para todo $i = 1, \dots, k$.

Definição 3 (Pareto ótimo próprio) $x^* \in S$ é Pareto ótimo próprio se ele é Pareto ótimo e se existe um $M > 0$ tal que para cada f_i e cada $x \in S$ satisfazendo $f_i(x) < f_i(x^*)$, existe no mínimo um f_j tal que $f_j(x) > f_j(x^*)$ e

$$\frac{f_i(x^*) - f_i(x)}{f_j(x) - f_j(x^*)} \leq M$$

para todo $i, j = 1, \dots, k$ e $i \neq j$. Uma solução que não é Pareto ótimo próprio é denominada solução Pareto ótimo impróprio.

Em problemas de otimização escalar, não há, em geral, soluções globais, e a condição usual para a existência de tais soluções é a convexidade. Em otimização multiobjetivo temos situação semelhante. Podemos ainda caracterizar a solução do POM localmente em uma vizinhança $B(x^*, r), r \geq 0, x^* \in S^*$, onde S^* é o conjunto Pareto ótimo.

Teorema 1 . *Se o POM é convexo então toda solução de Pareto ótimo local é também Pareto ótimo global.*

Prova 1 *Para prova veja teorema 2.2.3 em [2].*

c.q.d

2 Método ϵ -restrição

Em otimização multiobjetivo se utiliza métodos de escalarização durante o processo solução. Através de métodos de escalarização se transforma o POM em um problema de otimização escalar, o qual possui métodos de resolução bem desenvolvidos. Utilizaremos o método ϵ -restrição(P_ϵ) para resolver o POM.

No método ϵ -restrição um componente do vetor objetivo é selecionado para ser otimizado e todos os outros componentes objetivos são convertidos em restrições com a determinação de um limite superior para cada um deles. Temos que o POM transforma-se no seguinte problema,

$$\begin{aligned} \text{Problema 2 (P}\epsilon\text{)} \quad & \text{Minimizar} \quad f_l(x), \\ & \text{Sujeito a} \quad f_j(x) \leq \epsilon_j, \quad j \in \{1, \dots, k\}, j \neq l, \\ & \quad \quad \quad x \in S \subset \mathbb{R}^n. \end{aligned}$$

onde $l \in \{1, \dots, k\}$, $S_l(\epsilon) = \{x \in S \mid f_j(x) \leq \epsilon_j, j \neq l\}$ e $\epsilon \in \epsilon = \{(\epsilon_1, \dots, \epsilon_k)^T \in \mathbb{R}^k \mid S_l(\epsilon) \neq \emptyset\}$. Para x ser viável ao problema P_ϵ é necessário que $x \in S \cap S_l(\epsilon)$

Transformando o problema do exemplo 1 podemos ter o seguinte P_ϵ ,

$$\begin{aligned} \text{Exemplo 2 } [P_1(\epsilon)] \quad & \text{minimizar} \quad (x_1 - 3)^2 + (x_2 - 2)^2 \\ & \text{sujeito a} \quad x_1 + x_2 \leq \epsilon_2 \\ & \quad \quad \quad x_1 + 2x_2 \leq \epsilon_3 \\ & \quad \quad \quad -x_1 \leq 0 \\ & \quad \quad \quad -x_2 \leq 0 \end{aligned}$$

Graficamente temos que a função objetivo do $P_1(\epsilon)$ define um círculo de centro no ponto $(3, 2)$. Resolver esse problema é equivalente a se encontra um ponto viável o mais próximo do ponto $(3, 2)$. De acordo com região viável determinada pelos limitantes ϵ_2 e ϵ_3 tem-se que x^* é uma solução do exemplo 2.

Observe que o conjunto Pareto ótimo é determinado por pontos no interior e na fronteira do triângulo definido pelos pontos extremos $(1, 0)$, $(2, 0)$ e $(3, 2)$ e o segmento entre $(0, 0)$ e $(1, 0)$.

Em [1] e [2] podemos encontrar a caracterização da relação entre as soluções do P_ϵ e as soluções Pareto do problema de otimização multiobjetivo através de teoremas.

Uma grande dificuldade na formulação do P_ϵ é a determinação do ϵ . Uma alternativa para determinação de ϵ adequado é o uso do conceito de solução utópica.

Definição 4 (Solução utópica) *Uma solução utópica $f^{**} = (f_1^{**}, f_2^{**}, \dots, f_k^{**})$ do POM é definida como*

$$f_i^{**} = f_i(x^{*i}), \quad i = 1, \dots, k$$

onde,

$$x^{*i} = \arg \min_{x \in S} f_i(x)$$

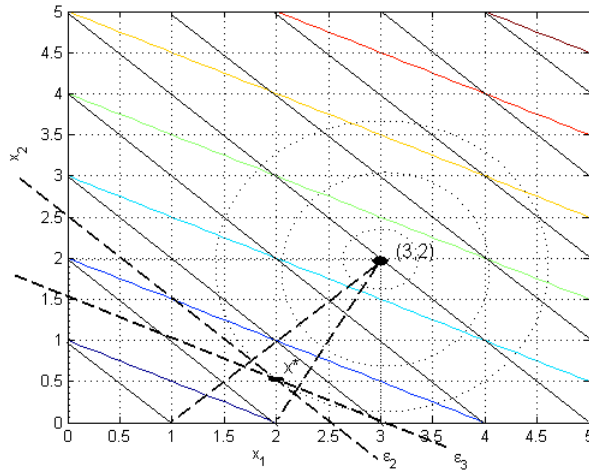


Figura 1: Gráfico do $P_1(\epsilon)$

É necessário encontrar valores adequados a inicialização do vetor $\epsilon = (\epsilon_1, \dots, \epsilon_k)$. Resolvendo os k problemas mono-objetivos podemos obter valores ótimos individuais para cada objetivo f_i , $i = 1, \dots, k$, os quais compõem o objetivo f^{**} correspondente à solução utópica do problema.

Paralelamente quando f_i está no seu ótimo individual podemos determinar os piores valores atingidos por cada objetivo f_j , $j = 1, \dots, k$, $j \neq i$, formando o vetor $f^0 = (f_1, f_2, \dots, f_k)$.

Em cada $P\epsilon$ os limites ϵ_i das restrições referente a f_i , $i = 1, \dots, k$ são determinados através de um gerador de números aleatórios com distribuição de probabilidades uniforme atendendo à restrição $f^{**} \leq \epsilon \leq f^0$.

A geração dos ϵ_j não é tarefa fácil, pois o ponto obtido pode não ser Pareto ótimo assim como os valores de ϵ_j podem tornar o problema inviável.

3 Taxas trade-off e os multiplicadores de Karush-Kuhn-Tucker

Definição 5 . Sejam $x^0, x^* \in S^*$ e as correspondentes imagens dos objetivos $f(x^0) = (f_1(x^0), \dots, f_k(x^0))$ e $f(x^*) = (f_1(x^*), \dots, f_k(x^*))$. Taxa trade-off entre x^0 e x^* envolvendo os objetivos f_i e f_j é definida como

$$T_{ij}(x^0, x^*) = \frac{f_i(x^0) - f_i(x^*)}{f_j(x^0) - f_j(x^*)}$$

onde $f_j(x^0) \neq f_j(x^*)$.

É possível estender a definição 5 de trade-off para o caso onde $f_j(x^0) = f_j(x^*)$, para maiores detalhes veja por exemplo [1].

Definição 6 . $T_{ij}(x^0, x^*)$ é chamada trade-off parcial envolvendo f_i e f_j entre x^0 e x^* se $f_l(x^0) = f_l(x^*)$ para todo $l = 1, \dots, k$ e $l \neq j, i$. Quando $f_l(x^0) \neq f_l(x^*)$ para no mínimo um $l \in \{1, \dots, k\}$, $T_{ij}(x^0, x^*)$ é chamada trade-off total envolvendo f_i e f_j entre x^0 e x^* .

Definição 7 . Dado $x^* \in S^*$. Suponha que exista uma direção viável d . Se o limite

$$\begin{aligned} t_{ij}(x^*, d) &= \lim_{\alpha \searrow 0} T_{ij}(x^* + \alpha d, x^*) \\ &= \lim_{\alpha \searrow 0} \frac{f_i(x^* + \alpha d) - f_i(x^*)}{f_j(x^* + \alpha d) - f_j(x^*)} \end{aligned}$$

existe, denominamos $t_{ij}(x^*, d)$ taxa trade-off em x^* na direção d envolvendo os objetivos f_i e f_j .

Se f_i e f_j são ambas de classe C^1 em x^* temos,

$$t_{ij}(x^*, d) = \frac{\nabla f_i(x^*)^T d}{\nabla f_j(x^*)^T d}$$

Também é possível estender a definição 7 de trade-off para o caso onde $f_j(x^* + \alpha d) = f_j(x^*)$, para maiores detalhes veja por exemplo [1].

O conceito chave entre os multiplicadores de Karush-Kuhn-Tucker λ_{kj}^* e taxa *trade-off* entre f_k e f_j em $x^* \in S^*$ é a interpretação da sensibilidade entre os multiplicadores. Para detalhes veja Teorema da Sensibilidade ([4] e [3]).

Teorema 2 . Suponha que x^* resolve $P\epsilon$ para algum $\epsilon^* \in \epsilon$, com

1. x^* sendo um ponto regular com respeito às restrições ativas do $P\epsilon$,
2. as condições suficientes de segunda-ordem satisfeitas em x^* , e
3. todas as restrições ativas em x^* não-degeneradas.

Assuma sem perda de generalidade que $\lambda_{kj}^* > 0$ para todo $j = 1, \dots, p$ e $\lambda_{kj}^* = 0$ para todo $j = p+1, \dots, k-1$. Então,

a). Se $\epsilon_j^* = f_j(x^*)$ para todo $j = 1, \dots, k-1$, então existe uma vizinhança $B(x^*; r)$ de x^* , $r > 0$ e uma função $\bar{x}(\cdot) : \mathbb{R}^{k-1} \rightarrow \mathbb{R}^n$ de classe C^1 definida em alguma vizinhança $B(\epsilon^*; R) \subset \mathbb{R}^{k-1}$ de ϵ^* , $R > 0$ tal que

$$S^* \cap B(x^*; r) \subseteq \bar{x}(B(\epsilon^*; R)) \subseteq S^* \quad (1)$$

onde S^* é o conjunto Pareto ótimo no espaço decisão.

b). Suponha $p = k-1$ e $B(x^*; r)$ são obtidos como em a). Sejam

$$Z_k^* = \{(f_1, \dots, f_k)^T \mid f_j = f_j(x), j = 1, \dots, k, x \in S^* \cap B(x^*; r)\}$$

e

$$Z_{k-1}^* = \{(f_1, \dots, f_{k-1})^T \mid f_j = f_j(x), j = 1, \dots, k-1, x \in S^* \cap B(x^*; r)\}.$$

Então existe uma função \bar{f}_k de classe C^1 definida em Z_{k-1}^* tal que para cada

$$(f_1, \dots, f_k)^T \in Z_k^*, f_k = \bar{f}_k(f_1, \dots, f_{k-1})^T.$$

Além disso

$$\frac{\partial \bar{f}_k}{\partial f_j}(f_1(x^*), \dots, f_{k-1}(x^*)) = -\lambda_{kj}^* \quad (2)$$

para cada $j = 1, \dots, k-1$.

c). Seja $1 \leq p \leq k-1$, onde $Z_\epsilon^* = \{(f_1, \dots, f_k)^T \mid f_j = f_j(x), j = 1, \dots, k, x \in \bar{x}(B(\epsilon^*; R))\}$. Então existem funções de classe C^1 , $\bar{f}_{p+1}, \dots, \bar{f}_{k-1}$ e \bar{f}_k definidas em $B(\epsilon^*; R)$ tais que para cada $(f_1, \dots, f_k)^T \in Z_\epsilon^*$ temos,

$$\bar{f}_j = \bar{f}_j(f_1, \dots, f_p, \epsilon_{p+1}^*, \dots, \epsilon_{k-1}^*), \text{ para todo } j = p+1, \dots, k.$$

Além disso para cada $i = 1, \dots, p$ temos

$$\frac{\partial \bar{f}_k}{\partial f_i} \Big|_{\epsilon = \epsilon^*} = -\lambda_{ki}^* = \frac{\nabla f_k(x^*)^T d_i^*}{\nabla f_i(x^*)^T d_i^*}$$

onde d^* é a direção $\frac{\partial \bar{x}(\epsilon^*)}{\partial \epsilon_l}$. Também temos que para cada $j = p+1, \dots, k-1$,

$$\frac{\partial \bar{f}_j}{\partial f_i} \Big|_{\epsilon = \epsilon^*} = \nabla f_j(x^*)^T \frac{\partial \bar{x}(\epsilon^*)}{\partial \epsilon_l} \quad (3)$$

Prova 2 . Para prova veja teorema 4.30 em [3].

c.q.d

Aplicando o teorema 2 ao exemplo 2.

Exemplo 3 Considere o problema $P_1(\epsilon)$ do exemplo 2 onde $\epsilon = (\epsilon_2, \epsilon_3)$. Seja $\epsilon_2 = 2,5$ e $\epsilon_3 = 3$. O ponto $x^0 = (2; 0,5)$ é solução do problema, já que x^0 é um ponto regular, os gradientes das restrições ativas $\nabla f_2(x^0) = (1, 1)^T$ e $\nabla f_3(x^0) = (1, 2)^T$ são linearmente independentes e as condições de Karush-Kuhn-Tucker são satisfeitas (Figura 2).

Em particular,

$$\nabla f_1(x^0) + \lambda_{12}^0 \nabla f_2(x^0) + \lambda_{13}^0 \nabla f_3(x^0) = 0$$

o que implica que $\lambda_{12}^0 = 1$ e $\lambda_{13}^0 = 1$.

Para este exemplo temos a seguinte função lagrangeana,

$$L(x, \lambda, \mu) = (x_1 - 3)^2 + (x_2 - 2)^2 + \lambda_{12}(x_1 + x_2 - \epsilon_2) + \lambda_{13}(x_1 + 2x_2 - \epsilon_3) - \mu_1 x_1 - \mu_2 x_2.$$

Calculando gradiente e hessiana,

$$\nabla L(x, \lambda, \mu) = (2(x_1 - 3) + \lambda_{12} + \lambda_{13} - \mu_1, 2(x_2 - 2) + \lambda_{12} + 2\lambda_{13} - \mu_2),$$

e

$$H(x, \lambda, \mu) = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}.$$

No ponto $x^0 = (2; 0,5)$ temos que $\nabla L(x^0, \lambda, \mu) = (0, 0)$ e $H(x^0, \lambda, \mu)$ é definida positiva.

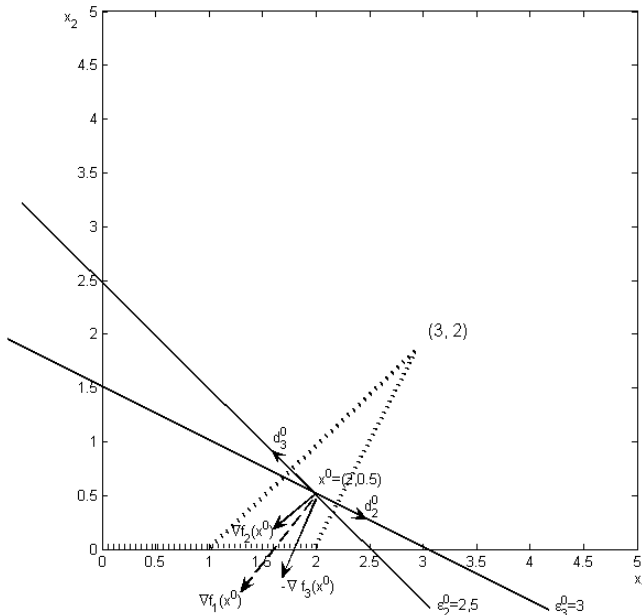


Figura 2: $P_1(2,5;3)$

É possível obter vários valores para $\epsilon = (\epsilon_1, \epsilon_2)$ de forma que ainda assim conseguimos soluções Pareto ótimo para $P_1(\epsilon)$. Temos que para $B(\epsilon^0; R)$ onde $R > 0$, existe uma função $\bar{x} : B(\epsilon^0; R) \rightarrow \mathbb{R}^k$ definida pela regra: para cada $\epsilon \in B(\epsilon^0; R)$, ϵ resolve $P_1(\epsilon)$ (Teorema 2).

Após encontrar essa vizinhança calculamos $\bar{x}(\epsilon)$, $d_2^0 = \frac{\partial \bar{x}(\epsilon^0)}{\partial \epsilon_2}$ e $d_3^0 = \frac{\partial \bar{x}(\epsilon^0)}{\partial \epsilon_3}$.

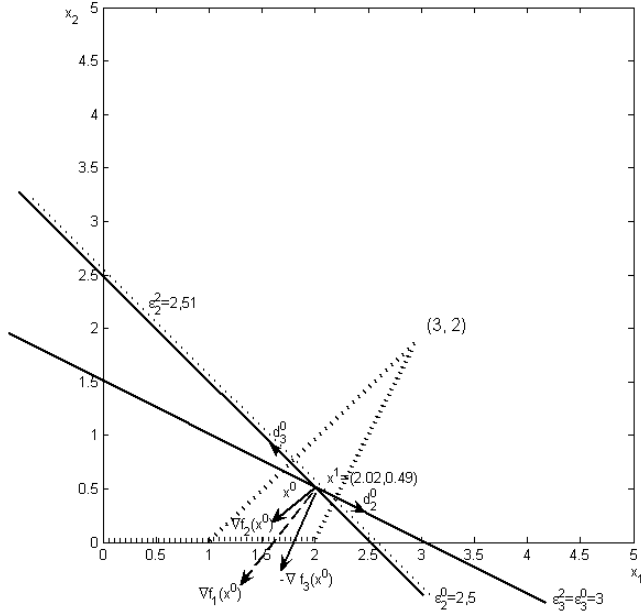


Figura 3: $P_1(2, 51; 3)$

Para algum $\epsilon \in B(\epsilon^0; R)$, a solução única de $P_1(\epsilon)$ é sempre um ponto na interseção das duas restrições, $f_2(x) \leq \epsilon_2$ e $f_3(x) \leq \epsilon_3$ (Figura 2). Assim para algum $\epsilon \in B(\epsilon^0; R)$, temos que $\bar{x}(\epsilon)$ é dado por,

$$f_2(x) = x_1 + x_2 = \epsilon_2, \quad f_3(x) = x_1 + 2x_2 = \epsilon_3. \quad (4)$$

Calculando $\bar{x}(\epsilon)$ usando 4,

$$\bar{x}(\epsilon) = \begin{bmatrix} \bar{x}_1(\epsilon) \\ \bar{x}_2(\epsilon) \end{bmatrix} = \begin{bmatrix} 2\epsilon_2 - \epsilon_3 \\ \epsilon_3 - \epsilon_2 \end{bmatrix}. \quad (5)$$

Calculando d_2^0 e d_3^0 (Figura 2) derivando 5,

$$d_2^0 = \frac{\partial \bar{x}(\epsilon^0)}{\partial \epsilon_2} = \begin{bmatrix} 2 \\ -1 \end{bmatrix}, \quad d_3^0 = \frac{\partial \bar{x}(\epsilon^0)}{\partial \epsilon_3} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}.$$

Em seguida examinamos $f_1(f_2, f_3)$. Para todo $\epsilon \in B(\epsilon^0)$ as desigualdades $f_2 \leq \epsilon_2$ e $f_3 \leq \epsilon_3$ são sempre ativas na solução ótima $\bar{x}(\epsilon)$ de $P_1(\epsilon)$. Para cada $\epsilon \in B(\epsilon^0; R)$ no caso de $P_1(\epsilon_2^0, \epsilon_3^0)$,

$$f_1(\bar{x}(\epsilon)) = [(2\epsilon_2^0 - \epsilon_3^0) - 3]^2 + [(\epsilon_3^0 - \epsilon_2^0) - 3]^2, \quad f_1(f_2, f_3) = [(2f_2 - f_3) - 3]^2 + [(f_3 - f_2) - 3]^2.$$

Segue que,

$$\frac{\partial f_1(x^0)}{\partial f_2} = 2(2f_2 - f_3 - 3)2 - 2(f_3 - f_2 - 2) = -1 = -\lambda_{12}^0.$$

e

$$\frac{\partial f_1(x^0)}{\partial f_3} = -2(2f_2 - f_3 - 3) - 2(f_3 - f_2 - 2) = -1 = -\lambda_{13}^0$$

O resultado acima possibilita mudarmos de ponto ao invés de ficar com o ponto x^0 . Podemos decidir, por exemplo, permanecer no nível de f_3 e alterar o nível de f_2 , por exemplo, aumentar

f_2 por no máximo $\delta\epsilon_2$, onde $0 < \delta\epsilon_2 \leq 1$. De acordo com essa decisão a melhor alternativa seria $x^1 = \bar{x}(\epsilon_2 + \delta\epsilon_2, \epsilon_3)$, onde

$$f_2(x^1) = \epsilon_2^0 + \delta\epsilon_2 f_3(x^1) = \epsilon_3^0.$$

e

$$f_1(x^1) = [2f_2(x^0) - f_3(x^0) - 3]^2 + [f_3(x^0) - f_2(x^0) - 2]^2 \approx f_1(x^0) + \lambda_{12}\delta\epsilon_2$$

Movendo de x^0 para x^1 ocorre um movimento na direção d_2^0 no espaço decisão (Figura 2). Por exemplo, se $\delta\epsilon_2 = 0,01$, a nova solução Pareto ótimo é (Figura 3),

$$\begin{bmatrix} \bar{x}_1^2(\epsilon) \\ \bar{x}_2^2(\epsilon) \end{bmatrix} = \begin{bmatrix} 2\epsilon_2^2 - \epsilon_3^2 \\ \epsilon_3^2 - \epsilon_2^2 \end{bmatrix} = \begin{bmatrix} 2,02 \\ 0,49 \end{bmatrix}$$

com $f_2(x^1) = 2,51$, $f_3(x^1) = 3$ e

$$f_1(x^2) = [2f_2(x^2) - f_3(x^2) - 3]^2 + [f_3(x^2) - f_2(x^2) - 2]^2 = 3,2405$$

Assim a mudança de $f_1(x^0) = 3,25$ para $f_1(x^1) = 3,2405$ a qual é igual a $f_1(x^1) - f_1(x^0) = -0,0095$ unidades é dada por $-\lambda_{12}^0\delta\epsilon_2 \approx -0,01$ unidades, justificando o uso de $-\lambda_{12}^0$ como uma taxa trade-off aproximada em x^0 .

4 Conclusão

Em virtude do que foi mencionado é possível se resolver problemas de otimização multiobjetivo através de métodos de otimização escalar, por exemplo o método $P\epsilon$. Após resolver o problema utiliza-se informações *trade-off* para decidir qual a melhor solução Pareto ótimo.

Percebe-se que os multiplicadores de Karush-Kuhn-Tucker, ferramentas matemáticas de grande utilidade em otimização são utilizados nesse processo. Através da análise desses multiplicadores se chega a uma solução desejável. Os multiplicadores podem ser interpretados como taxas, as quais determinam o ganho em uma determinada variável objetivo com uma perda em outra, enquanto as outras permanecem inalteradas. Os multiplicadores de Karush-Kuhn-Tucker são interpretados nesse caso como informações *trade-off*.

Referências

- [1] Yacov H. Haimes and Vira Chankong, Kuhn-Tucker Multipliers as Trade-Offs in Multiobjective Decision-Making Analysis, *Automatica*, 15 (1979), 59-72.
- [2] Kaisa Miettinen, "Nonlinear Multiobjective Optimization", International Series in Operations Research and Management Science, 12 (1999), Kluwer Academic Publishers, Dordrecht.
- [3] Yacov H. Haimes and Vira Chankong, "Multiobjective decision making: theory and methodology", Dover, Mineola, USA, (1983).
- [4] David E. Luenberger, "Linear and Nonlinear Programming", Second Edition, Addison-Wesley, Reading, MA, USA, 1984.
- [5] Hwang, C.L. and Masud, A.S., Multiple Objective Decision Making - Methods and Applications: A State-of-the-Art Survey, Lecture Notes in Economics and Mathematical Systems 164, Springer-Verlag, Berlin, Heiderberg, 1979.
- [6] Paulo Augusto Valente Ferreira, "Otimização Multiobjetivo: Teoria e Aplicações", Tese de Doutorado, FEEC-UNICAMP, 1998.