



O Problema de Recobrimento de Sólidos por Esferas de Diferentes Raios

Renan Vicente Pinto

Universidade Federal do Rio de Janeiro

renanvp@cos.ufrj.br

Nelson Maculan

Universidade Federal do Rio de Janeiro

maculan@cos.ufrj.br

RESUMO

Neste trabalho, apresentamos um modelo de programação matemática para o problema de recobrimento de sólidos por esferas de diferentes raios. Dado um conjunto de esferas, possivelmente de diferentes diâmetros, e um sólido, deseja-se posicionar essas esferas de tal modo que a união delas forme uma cobertura para esse sólido, utilizando a menor quantidade possível de esferas desse conjunto. Esse problema tem aplicação no planejamento do tratamento por radiocirurgia conhecida como *Gamma Knife* e pode ser formulado como um problema de otimização não-convexa, com restrições quadráticas e função objetivo linear.

PALAVRAS CHAVE. Problema de Recobrimento, Otimização Não Convexa, Otimização Combinatória.

Área Principal: Otimização Combinatória, Otimização Não Convexa

ABSTRACT

In this work, we present a mathematical programming model for the problem of covering solids by spheres of different radii. Given a set of spheres, possibly with different diameters, and a solid, the goal is to locate the spheres in such a way their union forms a coverage for this solid, using the smallest possible number of spheres of this set. This problem has an application in the radiosurgical treatment planning known as *Gamma Knife* and can be formulated as a nonconvex optimization problem with quadratic constraints and a linear objective function.

KEYWORDS. Covering Problem. Nonconvex Optimization. Combinatorial Optimization.

Main Area: Combinatorial Optimization, Nonconvex Optimization

1. O Problema de Recobrimento

Podemos descrever matematicamente o problema da seguinte forma:

(P) Dados um conjunto compacto $T \subset \mathbb{R}^3$, um conjunto finito de raios $R \subset \mathbb{R}_+$, um conjunto N indexando as esferas e uma função $\rho : N \rightarrow R$, encontrar um conjunto de esferas $\{B(x(i), \rho(i)) \mid i \in N\}$ de cardinalidade mínima recobrendo cada ponto de T .

Em outras palavras, dado um volume T a ser recoberto, devemos posicionar a menor quantidade possível de esferas de forma a realizar o recobrimento de T . Por recobrimento, entende-se que cada ponto de T deve pertencer, também, a alguma esfera. O problema (P), como definido acima, pode ser formulado como um problema de programação matemática, com o objetivo de encontrar os centros de cada esfera e associar a cada uma delas um respectivo raio, através da função ρ , de forma que a união dessas esferas forme uma cobertura do conjunto T .

Nas Figuras 1 e 2, são exemplificados o caso em que as esferas recobrem totalmente o volume T e o caso em que não há recobrimento total, respectivamente.

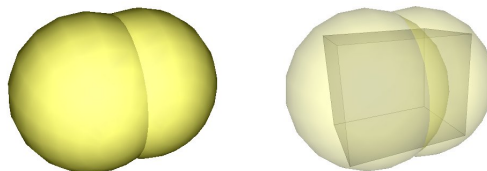


Figura 1: Exemplo de recobrimento para T sendo um retângulo. Como T está totalmente recoberto, não é visível na imagem à esquerda. À direita, é utilizado o efeito de transparência para melhor visualização.

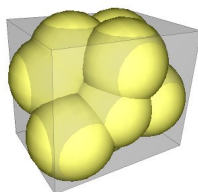


Figura 2: Exemplo onde não há recobrimento total.

2. Formulação Matemática

Um modelo de programação infinita não-linear inteira mista foi desenvolvido por Maculan (2009), com um número infinito não-enumerável de variáveis e restrições, como segue.

Sejam N , R e T como descritos anteriormente em (P). Sejam, ainda, U o conjunto de índices para os elementos de R , uma constante $M \geq \text{diam}(T)$, suficientemente grande, e uma constante $\epsilon > 0$, com $|\epsilon| \ll 1$. Considere as seguintes variáveis:

- $x^i : N \rightarrow \mathbb{R}^3$, que representa o centro da esfera i ;
- $y_i : N \rightarrow \{0, 1\}$, que assume valor $\begin{cases} 1, & \text{se a esfera } S_i \text{ é usada na solução,} \\ 0, & \text{caso contrário;} \end{cases}$
- $u_i(p) : N \times T \rightarrow \{0, 1\}$, que assume valor $\begin{cases} 1, & \text{caso a esfera } S_i \text{ cubra o ponto } p \in T, \\ 0, & \text{caso contrário;} \end{cases}$

- $w_{ij} : N \times U \rightarrow \{0, 1\}$, que assume valor $\begin{cases} 1, & \text{caso a esfera } S_i \text{ tenha o raio } r_j \in R, \\ 0, & \text{caso contrário;} \end{cases}$

O problema é corretamente descrito pelo modelo (1)-(6):

$$\min \sum_{i=1}^n y_i \quad (1)$$

$$\|x^i - p\|^2 \leq u_i(p) \sum_{j \in U} w_{ij} r_j^2 + (1 - u_i(p)) M^2, \quad \forall i \in N, \forall p \in T \quad (2)$$

$$\sum_{j \in U} w_{ij} = 1, \quad \forall i \in N \quad (3)$$

$$\sum_{i \in N} u_i(p) \geq 1, \quad \forall p \in T \quad (4)$$

$$\int_{p \in T} u_i(p) dp \geq \epsilon y_i, \quad \forall i \in N \quad (5)$$

$$\int_{p \in T} u_i(p) dp \leq \text{Vol}(T) y_i, \quad \forall i \in N \quad (6)$$

As restrições (2) fazem com que, se a esfera S_i cobre o ponto $p \in T$, então a distância euclidiana entre p e o centro x^i dessa esfera deve ser, no máximo, o raio dessa esfera. O fato de a cada esfera poder ser atribuído apenas um raio é representado pelas restrições (3). As restrições (4) asseguram que cada ponto $p \in T$ será recoberto por, pelo menos, uma esfera. Se a esfera S_i foi selecionada para a solução, as restrições (5) asseguram que ela recobre um volume não nulo de T . Caso contrário, as restrições (6) forçam as funções u_i a serem funções nulas.

Repare que, em (2) e em (4), temos uma restrição para cada ponto $p \in T$, resultando em um número infinito, não-enumerável, de restrições.

3. Modelo Proposto

Proporemos um modelo para o problema de recobrimento de uma região tridimensional por esferas de raios diferentes, como em Pinto (2011). Começaremos com algumas definições.

Seja S um conjunto de esferas, possivelmente de raios diferentes. Conforme (P), que denotaremos de problema *clássico*, para futuras referências no texto, o problema consiste em recobrir um sólido $T \subset \mathbb{R}^3$ com o menor número possível de esferas pertencentes a S . Ou, de outra forma, deve-se posicionar a menor quantidade possível de esferas de S de modo que cada ponto $p \in T$ pertença, também, a alguma dessas esferas. Propomos, então, uma descrição formal do problema, em termos de uma formulação de programação matemática.

Seja S um conjunto formado por n esferas, $S = \{S_1, \dots, S_n\}$. Assim, para cada $i \in \{1, \dots, n\}$, seja S_i uma esfera centrada em $x^i \in \mathbb{R}^3$ e com raio $r_i > 0$. Assumiremos que um recobrimento de T por esferas de S existe. Ou, de outra forma, assumiremos que há disponível, em S , esferas suficientes capazes de recobrir T . Temos, então, as seguintes variáveis:

- $x \in T^n$, e

- $y \in \{0, 1\}^n$, sendo $y_i = \begin{cases} 1, & \text{se a esfera } S_i \text{ é utilizada} \\ 0, & \text{caso contrário} \end{cases}, \quad \forall i \in \{1, \dots, n\}.$

A partir da idéia de dualidade em programação matemática, montaremos um modelo bastante parecido com um modelo de empacotamento, mas com as restrições de não-sobreposição

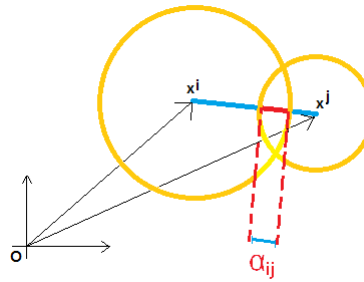


Figura 3: Representação dos parâmetros α , que correspondem à máxima interpenetração das esferas.

relaxadas, de forma a possibilitar um recobrimento. Tal relaxamento é controlado pelos parâmetros α_{ij} , com $1 \leq i < j \leq n$, conforme representado na Figura 3.

Ao invés de forçarmos que a distância entre os centros de duas esferas dadas deve ser maior que a soma de seus raios, permitiremos que essas esferas se sobreponham a uma quantidade máxima α_{ij} .

Para duas esferas S_i e S_j , a distância mínima entre seus centros é, então, dada por:

$$d_{ij} = r_i + r_j - \alpha_{ij} \quad (7)$$

Assim, as restrições do modelo serão:

$$\|x^i - x^j\| \geq d_{ij}, \quad \forall 1 \leq i < j \leq n. \quad (8)$$

Porém, como a função norma euclideana apresenta pontos de não-diferenciabilidade, suavizaremos as restrições elevando ambos os lados da igualdade ao quadrado. Ainda, essas restrições somente são necessárias se ambas as esferas S_i e S_j estiverem presentes na solução. Logo, substituiremos as restrições 8 por:

$$\|x^i - x^j\|^2 \geq d_{ij}^2 (y_i + y_j - 1), \quad \forall 1 \leq i < j \leq n. \quad (9)$$

No lado direito da equação 9, introduzimos o termo $(y_i + y_j - 1)$, que só assume valor 1, recuperando a equação original 8, quando $y_i = 1$ e $y_j = 1$, ou seja, quando ambas as esferas S_i e S_j estão presentes na solução. Se somente uma das esferas estiver presente na solução, esse termo assumirá valor 0 e, caso nenhuma delas esteja presente, assumirá valor -1. Nesses dois casos, teremos redundância, já que a norma assume sempre um valor positivo.

Por fim, tomemos os parâmetros c_i como sendo o custo de cada esfera na solução. Como temos por objetivo a menor quantidade possível de esferas na solução, priorizaremos as esferas de maior raio, atribuindo a elas um custo maior, visto que a formulação será um problema de maximização.

Propomos, então, a seguinte formulação:

$$\max \sum_{i=1}^n c_i y_i \quad (10)$$

$$\text{s. a } \|x^i - x^j\|_2^2 \geq (r_i + r_j - \alpha_{ij})^2 (y_i + y_j - 1), \quad \forall 1 \leq i < j \leq n \quad (11)$$

$$x \in T^n \quad (12)$$

$$y \in \{0, 1\}^n \quad (13)$$

A partir dessa formulação, escrevemos o seguinte lema, que define a existência de um conjunto de valores para os parâmetros, com os quais podemos retomar a solução do problema clássico:

Lema. Existe um conjunto de parâmetros $\alpha_{ij} \geq 0$, com $1 \leq i < j \leq n$, e um conjunto de valores $c_i \geq 0$, com $1 \leq i \leq n$, para os quais uma solução ótima de (10)-(13) é, também, uma solução ótima para o problema clássico de recobrimento (P).

Demonstração: Seja (x^*, y^*) uma solução ótima do problema clássico de recobrimento. Para cada par de esferas da solução, calculemos a distância D_{ij} entre seus centros:

$$D_{ij} = \|(x^i)^* - (x^j)^*\|, \quad \forall 1 \leq i < j \leq n.$$

Se $D_{ij} > r_i + r_j$, tome $\alpha_{ij} = 0$. Caso contrário, tome

$$\alpha_{ij} = r_i + r_j - D_{ij}.$$

Tome $c_i = y_i^*$, para todo $1 \leq i \leq n$. Então, afirmamos que (x^*, y^*) é uma solução ótima do problema (10)-(13), com os parâmetros $\{\alpha_{ij}\}_{1 \leq i < j \leq n}$ e $\{c_i\}_{1 \leq i \leq n}$ como acabamos de definir.

De fato, utilizando os parâmetros agora escolhidos, é fácil verificar que a solução (x^*, y^*) também será viável em (10)-(13). Basta mostrar, então, que ela terá o melhor valor da função objetivo. De fato, como os y_i^* assumem valores 0 ou 1, para todo i , e estamos usando $c_i = y_i^*$, temos que

$$\sum_{i=1}^n c_i y_i = \sum_{i|y_i^*=1} y_i.$$

Podemos observar que esse somatório atingirá seu valor máximo se tomarmos $y_i = 1$ quando $y_i^* = 1$, ou seja, quando tivermos $y_i = y_i^*$. Portanto, (x^*, y^*) é uma solução ótima de (10)-(13) com os parâmetros tais como definidos. ■

O modelo apresentado é um modelo de programação matemática não-linear inteira mista (MINLP) cuja relaxação contínua é não-convexa, devido às restrições (11). Embora alguns métodos de solução existam, essa classe de problemas está entre as mais difíceis de serem resolvidas. O estado da arte de tais métodos [Tawarmalani (2002)] não está em um estágio onde instâncias diversas podem ser resolvidas efetivamente ou asseguradamente.

Vale ressaltar que o problema não é bem definido. Se há disponível uma esfera suficientemente grande, ou com diâmetro maior que o diâmetro de T , ela sozinha seria capaz de recobrir T . Assim, definiremos uma região onde as esferas devem estar inteiramente contidas, como pode ser visto na Figura 4. Nos referiremos a essa região como *região de segurança*.

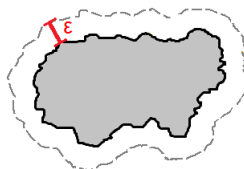


Figura 4: Representação da região de segurança. A região em cinza representa o volume a ser recoberto.

O domínio das variáveis x^i será restringido, então, de acordo com a região de segurança. Novamente, o ideal é o domínio da variável x^i ser o lugar geométrico dos pontos que distam r_i da região de segurança, de forma a manter a esfera S_i totalmente contida nessa região. Isso pode ser visto na Figura 5.

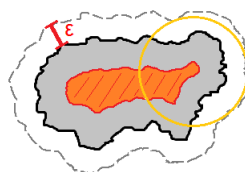


Figura 5: A região em cinza representa o volume a ser recoberto. Como as esferas devem estar contidas na região de segurança, o domínio da variável que representa o seu centro é restrito à região em laranja.

4. Resultados

Apresentaremos alguns resultados obtidos para o problema de recobrimento, utilizando o modelo proposto. Dentre os solver comerciais existentes, testamos o COUENNE, o BONMIN e o XPRESS. Ressaltamos que um recobrimento total é possível de ser atingido, desde que a região de segurança seja grande o suficiente e desde que as esferas possam ter a quantidade de interseção suficiente. Com a finalidade de visualização e melhoria manual das soluções, utilizamos um programa por nós desenvolvido, de onde foram retiradas todas as imagens que aparecem no restante deste trabalho.

Introduziremos três grandezas, como em Yun (2003), cuja finalidade é apenas conhecer melhor as características da solução. Seja T o volume a ser recoberto pelas esferas e U o volume da união das esferas. Calcularemos:

- ‘cov’ : Percentual do volume de T coberto pelas esferas;
- ‘overlap’ : Percentual do volume de T coberto por mais de uma esfera; e
- ‘miscov’ : Percentual do volume de U localizado na parte externa de T .

Essas grandezas podem ser melhor entendidas com o auxílio da Figura 6.

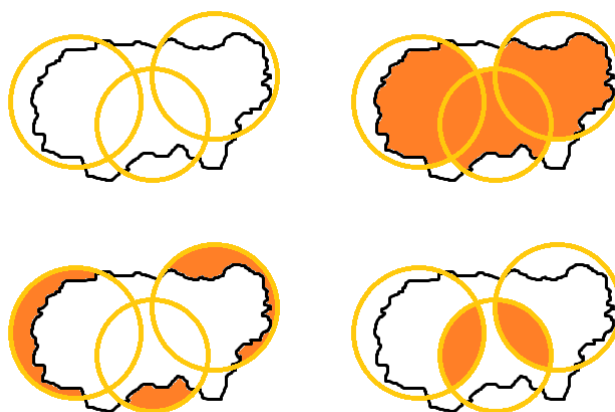


Figura 6: Exemplo bidimensional mostrando as regiões que correspondem às grandezas utilizadas na descrição de soluções. Em (a), a região a ser recoberta e as bolas presentes no recobrimento. As regiões hachuradas representam os parâmetros (b) cov, (c) miscov e (d) overlap.

O COUENNE (*Convex Over and Under ENvelopes for Nonlinear Estimation*) é um código open-source que consiste no algoritmo conhecido como *spatial Branch and Bound*. O pacote

implementa técnicas de linearização, técnicas de ramificação, heurísticas para encontrar soluções viáveis e técnicas de melhoria de limites inferiores para o problema.

Durante o teste, forçamos o término do algoritmo após 9 dias de execução. A melhor solução encontrada, depois de 20h de execução, está ilustrada na Figura 7. Para essa solução, temos $cov = 68,12\%$, $miscov = 7,66\%$ e $overlap = 9,03\%$. O valor dessa solução na função objetivo é 352.

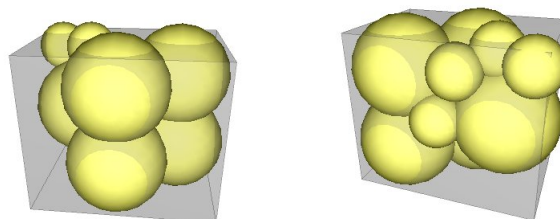


Figura 7: Imagens da solução encontrada pelo COUENNE, sendo disponibilizadas 6 esferas de raio 4mm e 20 esferas de raio 2mm.

O BONMIN (*Basic Open-source Nonlinear Mixed INteger programming*) implementa algoritmos exatos quando as funções são convexas. Mas são apenas heurísticas quando esse não é o caso. Para MINLPs não-convexos, o algoritmo padrão é um *Branch and Bound* baseado em programação não-linear. Para a solução encontrada, após 390 segundos, temos $cov = 84,08\%$, $miscov = 9,66\%$ e $overlap = 10,22\%$. O valor dessa solução é 448. Imagens da solução podem ser vistas na Figura 8.

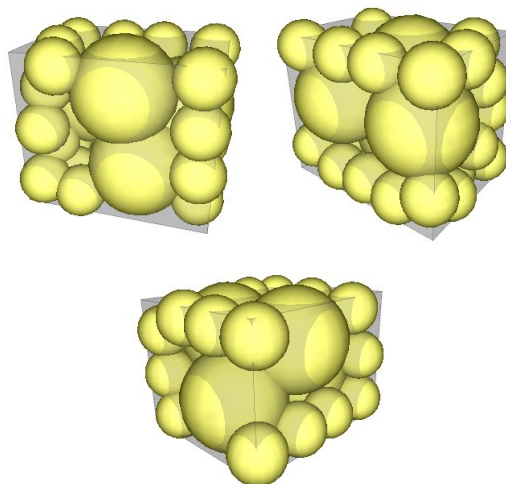


Figura 8: Imagens da solução encontrada pelo BONMIN, sendo disponibilizadas 6 esferas de raio 4mm e 24 esferas de raio 2mm.

O Xpress-SLP é um resolvidor para problemas de programação não-linear. Ele implementa o método de programação linear sequencial. Embora a solução seja o resultado de uma otimização da aproximação linear do problema, não há uma garantia de que ela será uma solução ótima do problema original não-linear. Ela pode ser apenas um ótimo local. Imagens da solução encontrada podem ser vistas na Figura 9.

O XPRESS-SLP apresentou desempenho melhor que o BONMIN, conseguindo uma solução de valor 496 na função objetivo, em apenas 4 segundos. Para essa solução, os valores dos parâmetros são $cov = 87,67\%$, $miscov = 10,58\%$ e $overlap = 15,87\%$.

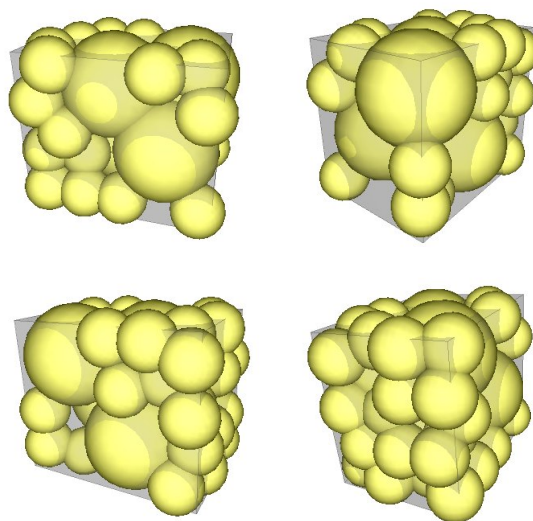


Figura 9: Imagens da solução encontrada pelo XPRESS-SLP, sendo disponibilizadas 6 esferas de raio 4mm e 30 esferas de raio 2mm.

5. Uma Aplicação

O problema de recobrimento por esferas de diferentes tamanhos tem uma importante aplicação no planejamento de tratamentos por radiocirurgias estereotáxicas, sendo uma das modalidades mais efetivas dessa classe de cirurgias a radiocirurgia conhecida como *Gamma Knife* [Ganz (1997)]. O *Gamma Knife* consiste em uma unidade de tratamento altamente especializada, direcionada a tratamentos de tumores cerebrais, malformações vasculares e distúrbios de dor na cabeça.

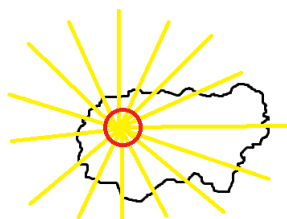


Figura 10: (a) Paciente na unidade de tratamento. (b) Diferentes tamanhos de elmos, resultando em 4 diferentes tamanhos de *shots*.

No interior da unidade de tratamento (ver Figura 10), feixes de 201 fontes radioativas de Cobalto-60 são irradiados e focados em um único ponto no espaço. Esse procedimento é conhecido como um *shot* de radiação. Um único raio não é suficiente para queimar consideravelmente o tecido. Logo, podemos considerar que a região afetada pelo *shot* é, aproximadamente, uma esfera centrada no foco de irradiação, como mostrado na Figura 11.

No equipamento afixado à cabeça do paciente, como mostrado na Figura 10, podem ser anexados 4 capacetes que se diferenciam no tamanho, como pode ser visto na Figura 10. Esses capacetes são responsáveis pelos tamanhos dos *shots*. Portanto, quatro tamanhos diferentes de *shots* estão disponíveis, afetando regiões aproximadamente esféricas com raios de 2mm, 4mm, 7mm e 9mm. Combinando múltiplos *shots* de radiação, um plano de tratamento pode ser realizado, de forma a tratar lesões de diferentes tamanhos e formas.

A região de segurança imposta ao modelo serve para proteger os tecidos saudáveis ao redor do tumor.

Figura 11: Representação de um *shot* de radiação.

6. Conclusão

Estudamos o problema de recobrimento de sólidos por esferas de diferentes raios e formulamos um modelo de programação matemática inteira mista não-linear (MINLP), cuja relaxação contínua é não-convexa. Constatamos que os resolvidores atuais para esse tipo de problema são bastante suscetíveis a ótimos locais. Realizamos testes com o modelo proposto e o COUENNE, o BONMIN e o Xpress-SLP. O COUENNE, apesar de implementar um algoritmo visando resolver problemas do tipo MINLP com relaxação contínuas não-convexas, conhecido como *Spatial Branch-and-Bound*, não foi capaz de obter bons resultados em tempos curtos. O BONMIN e o Xpress-SLP, para MINLPs não-convexos, são apenas heurísticas, já que implementam métodos baseados em aproximações do modelo original. Se mostraram bastante sensíveis aos dados, já que o simples acréscimo de uma esfera ao conjunto S , das esferas disponíveis, fez com que o solver retornasse que não há solução viável, em muitos casos. Dentre esses, o Xpress-SLP foi o que mostrou melhores resultados e em tempos bem curtos.

Como trabalhos futuros, pretendemos estudar reformulações do modelo proposto que permitam a obtenção de melhores soluções com o uso dos resolvidores atuais. Porém, o principal objetivo é um estudo de valores para os parâmetros ϵ , α e γ de forma que a solução do modelo seja um recobrimento completo, mas ainda matendo outros objetivos, como menor quantidade possível de esferas, com pouca sobreposição e pouca ocupação da região externa ao sólido a ser recoberto.

Referências

- Ganz, J. C.** (1997), Gamma Knife Surgery.
- Maculan, N. e Liberti, L. e Zhang, Y.** (2009), Optimal configuration of gamma ray machine radiosurgery units: the sphere covering subproblem, *Optimization Letters*, 3 (1), 109-121.
- Pinto, R. V.** (2011), Um Enfoque Para o Problema de Recobrimento de Sólidos por Esferas de Diâmetros Diferentes, *Dissertação de Mestrado, COPPE/UFRJ*.
- Tawarmalani, M. e Sahinidis, N.** (2002), Convexification and Global Optimization in Continuous and Mixed-Integer Nonlinear Programming: Theory, Algorithms, Software and Applications.
- Yun, L. e Yungqing, Y. e Zhen, W.** (2003), A Sphere Packing Model for the Optimal Treatment Plan, *The UMAP Journal*, 24 (3), 339-350.