# Análise do desempenho do aprendizado por reforço na solução do problema da mochila multidimensional

André Luiz Carvalho Ottoni <sup>1</sup> Erivelton Geraldo Nepomuceno <sup>2</sup> Marcos Santos de Oliveira <sup>3</sup>

**Resumo:** Neste trabalho, o objetivo é analisar o desempenho do Aprendizado por Reforço na solução do Problema da Mochila Multidimensional. Para isso, é proposto um modelo de Aprendizado por Reforço estruturado em estados, ações e recompensas. Além disso, os experimentos computacionais apresentados permitem analisar a sensibilidade dos parâmetros do algoritmo Q-learning na resolução desse tipo de problema de otimização combinatória.

**Palavras-chave:** Aprendizado por Reforço. Otimização Combinatória. Problema da Mochila Multidimensional.

**Abstract:** In this work, the goal is to analyze the performance of Reinforcement Learning in solving the Multidimensional Knapsack Problem. For this, a Reinforcement Learning model structured in states, actions and rewards is proposed. In addition, the computational experiments presented allow us to analyze the sensitivity of the parameters of the Q-learning algorithm in solving this type of combinatorial optimization problem.

**Keywords:** Combinatorial Optimization. Multidimensional Knapsack Problem. Reinforcement Learning.

## 1 Introdução

O Problema da Mochila (em inglês, *Knapsack Problem*) é um problema clássico de otimização combinatória [1] [2]. Nesse tipo de problema, o objetivo é encontrar uma combinação de itens que maximize o lucro de transporte desses objetos. A versão mais simples possui apenas uma mochila (unidimensional). Já o Problema da Mochila estruturado com *m* mochilas é mais complexo, e é denominado Problema da Mochila Multidimensional (em inglês, *Multidimensional Knapsack Problem*) [3] [2] [4].

Na literatura, o Problema da Mochila tem diversas aplicações como na otimização de corte de materias, problemas de embalagem e carregamento de veículos de transporte [1] [5] [4] [6]. Além disso, vários métodos já foram abordados na resolução do Problema da Mochila, como Algoritmos Genéticos [1] [7] [6], Algoritmos Evolucionários [8] [4], Programação Dinâmica [9], Busca Tabu [3] [10], Otimização por Colônia de Formigas [11] [12] e Otimização por Enxame de Partículas [13] [14]. No entanto, poucos trabalhos investigaram a aplicação do Aprendizado por Reforço (AR) para solucionar o Problema da Mochila. Alguns raros exemplos são os trabalhos de [15] e [16].

O AR é um campo da Aprendizagem de Máquina (em inglês, *Machine Learning*) que se baseia no aprendizado por interação em um ambiente [17] [18]. Assim, no AR um agente aprende a tomar decisões a partir dos sucessos e fracassos de suas ações [17]. Na literatura, o AR possui aplicações diversas, inclusive no campo da

http://dx.doi.org/10.5335/rbca.v9i3.6601

<sup>&</sup>lt;sup>1</sup>Departamento de Engenharia Mecatrônica, *Campus* Divinópolis, Centro Federal de Educação Tecnológica de Minas Gerais (CEFET-MG). {andreottoni@ymail.com}

<sup>&</sup>lt;sup>2</sup>Departamento de Engenharia Elétrica, Universidade Federal de São João del-Rei (UFSJ).

<sup>{</sup>nepomuceno@ufsj.edu.br}

<sup>&</sup>lt;sup>3</sup>Departamento de Matemática e Estatística, Universidade Federal de São João del-Rei (UFSJ). {mso@ufsj.edu.br}

otimização combinatória [19] [20] [21]. Nesse sentido, [22] e [21] realizam experimentos para analisar o desempenho do AR na solução do Problema do Caixeiro Viajante (PCV), de acordo com a definição dos parâmetros taxa de aprendizado  $(\alpha)$ , fator de desconto  $(\gamma)$  e política  $\epsilon-greedy$   $(\epsilon)$ . Os resultados de [22] e [21] apontam que a seleção dos valores de  $\alpha$ ,  $\gamma$  e  $\epsilon$  influenciam diretamente na busca por uma solução ótima no PCV. De fato, estudos já demonstraram que o desempenho do AR pode ser bem sensível a definição desses parâmetros [17] [23] [24] [25].

Dessa forma, o objetivo deste trabalho é analisar o desempenho do AR na solução do Problema da Mochila Multidimensional. Para isso, é proposto um modelo de AR estruturado em estados, ações e recompensas. Além disso, pretende-se analisar a sensibilidade dos parâmetros do AR na resolução desse tipo de problema de otimização combinatória.

Este trabalho está organizado em Seções. A Seção 2 apresenta conceitos teóricos iniciais do Problema da Mochila Multidimensional. Em seguida, as Seções 3 e 4 descrevem fundamentos teóricos do AR e o modelo proposto, respectivamente. Os experimentos realizados são apresentados na Seção 5. Já a Seção 6, descreve a análise dos resultados. Finalmente, na Seção 7 são apresentadas as conclusões.

#### 2 Problema da mochila multidimensional

No Problema da Mochila Multidimensional (PMM), o objetivo é selecionar um conjunto de objetos para serem alocados em m mochilas, de forma a maximizar o somatório dos valores dos itens, conforme a Equação (1). Para isso, deve ser respeitada a capacidade máxima de cada mochila (Equação (2)). Além disso, uma variável binária  $x_i$  armazena a decisão se um objeto i foi selecionado (1 = Sim ou 0 = Não) (3). Dessa forma, as Equações (1), (2) e (3) apresentam uma possível formulação matemática para o PMM [1] [2]:

Maximizar:

$$\sum_{i=1}^{n} v_i x_i. \tag{1}$$

Sujeito a:

$$\sum_{i=1}^{n} p_{ji} x_i \le c_j \qquad j = 1, ..., m.$$
 (2)

$$x_i \in \{0,1\} \qquad i = 1,...,n.$$
 (3)

Em que, n é o número total de itens disponíveis; m é o número de mochilas; i é o índice referente ao item; j é o índice referente à mochila;  $v_i$  é o valor do item i;  $x_i$  é a variável binária referente ao item i;  $p_{ji}$  é o peso do item i referente à mochila j; e  $c_j$  é a capacidade máxima da mochila j.

Assim, em termos matriciais o PMM pode ser estruturado com três matrizes V, P e C, em que, V é matriz de valores dos itens, de ordem  $1 \times n$ ; P é matriz de pesos dos itens para cada mochila, de ordem  $m \times n$ ; e C é a matriz que armazena as capacidades máximas de cada mochila, de ordem  $1 \times m$ .

No PMM, o espaço de busca para cada problema cresce exponencialmente com o número de itens, multiplicado pelo número de mochilas. Dessa forma, o número de soluções possíveis para cada instância é dado por  $m \times 2^n$  [4] [6].

## 3 Aprendizado por reforço

O Aprendizado por Reforço (AR) é uma técnica de Aprendizado de Máquina fundamentada nos Processos de Decisão de Markov, em inglês *Markov Decision Processes* (MDP) [17]. Um MDP modela processos em que as transições entre os estados do sistema são probabilísticas [26]. Além disso, são chamados de processos de decisão, pois modelam a possibilidade de um agente interferir no sistema executando ações [26]. O AR consiste em aprender mapeando ações para situações (estados), com o objetivo de maximizar ao longo do tempo o valor de recompensa [17].

Em seguida, uma breve descrição do algoritmo de AR adotado neste trabalho.

## 3.1 Algoritmo Q-learning

O *Q-learning* [27] é um dos algoritmos de AR mais tradicionais. O método (Algoritmo 1) se baseia na atualização da matriz de aprendizado Q, a partir da Equação (4):

$$Q_{t+1} = Q_t(s, a) + \alpha [r(s, a) + \gamma \max_a' Q(s', a') - Q_t(s, a)], \tag{4}$$

em que: Q é a matriz de aprendizado, r(s,a) é o reforço para a execução da ação a no estado s,  $\alpha$  é a taxa de aprendizado e  $\gamma$  é o fator de desconto.

```
1.
   Para cada (s,a) inicialize Q(s,a)=0;
2.
   Observe o estado s;
3.
   Repita até o critério de parada ser satisfeito
          Selecione a ação a usando a política \textit{e-greedy};
4.
          Execute a ação a;
5.
         Receba a recompensa imediata r(s,a);
6.
7.
         Observe o novo estado s';
         Atualize o item Q(s,a) de acordo com a Equação (1);
8.
9.
          s = s';
10. Fim Repita
```

## Algoritmo 1: Q-learning.

A matriz Q possui dimensões  $n_s \times n_a$ , em que,  $n_s$  é o número de estados do modelo e  $n_a$  é o número de ações que o agente pode realizar. Assim, em cada iteração o agente identifica a situação atual do ambiente (observa o estado), seleciona uma ação para esse estado e recebe uma reforço r(s,a), referente ao par estado s0 (a). Para a atualização de s0 também é necessário identificar o novo estado do ambiente s0 ou s1), já que, o termo s1, já que, o termo s2, s3 o máximo valor na linha relativa ao novo estado s3 na matriz.

# 3.2 Parâmetros

Nesta subseção são apresentados alguns parâmetros importantes para boa parte dos métodos de AR.

- Taxa de Aprendizado ( $\alpha$ ): controla a velocidade das atualizações.
- Fator de Desconto ( $\gamma$ ): controla o grau de influência de recompensas futuras no instante t sobre a recompensa imediata ( $r_{t+1}$ ).
  - $\gamma=1$ : as recompensas futuras são fortemente consideradas;
  - $-\gamma < 1$ : a influência do somatório das recompensas futuras é limitado;
  - $\gamma = 0$ : maximiza a recompensa imediatada, ou seja, o objetivo é simplesmente aprender uma ação  $(a_t)$  em um estado  $(s_t)$  para maximizar somente  $r_{t+1}$ .
- Parâmetro  $\epsilon$  da política  $\epsilon-greedy$ : responsável pelo controle entre gula e aleatoriedade na seleção das ações.

A regra de seleção de ações da política  $\epsilon-greedy$  é dada na Equação 5:

$$\pi(s) = \begin{cases} a^*, & \text{com probabilidade} \quad 1 - \epsilon \\ a_a, & \text{com probabilidade} \quad \epsilon, \end{cases}$$
 (5)

em que:

- $\pi(s)$  é a política de decisão para o estado corrente s;
- $a^*$  é a ação mais bem estimada para o estado s no instante atual;
- $a_a$  é uma ação aleatória selecionada com probabilidade  $\epsilon$ .

# 4 Modelo de aprendizado por reforço

O modelo de AR proposto para a resolução do Problema da Mochila é estruturado a partir da definição do conjunto de Ações, Estados e Recompensas:

- **Estados**: um estado  $(s_t)$  identifica qual objeto é inserido na mochila no instante t.
- Ações: uma ação (at) no instante t representa a intenção de inserir um determinado item na mochila em t + 1. Vale ressaltar que, ao colocar um objeto na mochila, o item é retirado da lista de ações disponíveis. Assim, ao finalizar um episódio de aprendizado, todos os objetos voltam a ficar disponíveis para reiniciar o processo.
- Recompensas: as recompensas buscam valorizar os objetos mais importantes da lista de itens, a partir da matriz V. Além disso, penalizar a inserção de itens com altos valores de restrição (pesos da matriz P). Assim, a recompensa (r(s,a)) ao inserir um item i na mochila j é dada pela Equação 6:

$$r(s,a) = v_i - p_{ji}, j = 1,...,m.$$
 (6)

Dessa forma, o modelo proposto tem como objetivo permitir ao agente aprender a sequência de itens que devem ser inseridos nas mochilas otimizando a função objetivo. Além disso, foi implementada uma rotina que impossibilita que o agente insira um objeto que ultrapassa as capacidades das m mochilas no instante t, garantindo assim, as restrições do modelo matemático do Problema da Mochila Multidimensional.

# 5 Experimentos realizados

Os experimentos compreenderam simulações no *software*  $MATLAB^{\textcircled{R}}$  com 12 instâncias da biblioteca ORLIB<sup>4</sup> [28], conforme apresentado na Tabela 1. O valor ótimo indica a melhor solução possível para o problema, ou seja, o valor que maximiza o somatório de itens na mochila, conforme Equação (1).

Para realizar uma análise de sensibilidade dos parâmetros do AR na solução do Problema da Mochila Multidimensional, foi abordada uma estrutura experimental baseada em trabalhos anteriores [22] [21] [25]. Assim, para cada instância foram realizadas simulações com 192 combinações dos parâmetros taxa de aprendizado  $(\alpha)$ , fator de desconto  $(\gamma)$  e parâmetro  $\epsilon$  da política  $\epsilon-qreedy$ :

- $\alpha$ ,  $\gamma$ : [0,01, 0,15, 0,30, 0,45, 0,60, 0,75, 0,90, 0,99],
- $\epsilon$ : [0, 0,05, 0,10].

Além disso, cada combinação foi simulada em cinco épocas (repetições) com 1000 episódios. Vale ressaltar que, a resposta de um episódio é o somatório dos valores dos itens adicionados nas m mochilas. Além disso, no início no primeiro episódio de uma época não existe aprendizado. Assim, o aprendizado vai sendo acumulado ao longo da época até o milésimo episódio (fim da época).

<sup>&</sup>lt;sup>4</sup><http://people.brunel.ac.uk/~mastjjb/jeb/orlib/files/mknap2.txt>

Tabela 1: Problemas da ORLIB estudados.

Tabela 1: 11001emas da Ortens estadados.							
	Problema	m	n	Ótimo			
1	sento1	30	60	7772			
2	weing1	2	28	141278			
3	weing2	2	28	130883			
4	weing3	2	28	95677			
5	weing7	2	105	1095445			
6	weish1	5	30	4554			
7	weish2	5	30	4536			
8	weish30	5	90	11191			
9	hp1	4	28	3418			
10	hp2	4	35	3186			
11	pb1	4	27	3090			
12	pb2	4	34	3186			

Fonte: elaborado pelos autores com base nos problemas da ORLIB estudados.

#### 6 Análise dos resultados

#### 6.1 Resultados gerais

Os resultados gerais dos experimentos realizados neste trabalho são descritos nesta seção. Nesse aspecto, são apresentadas as instâncias em que o AR conseguiu a solução ótima, e também uma solução não ótima. Além disso, são realizadas análises iniciais dos parâmetros do AR para as instâncias em que não se alcançou a solução ótima.

O AR alcançou a solução ótima em quatro instâncias: weing1, weing2, weing3. e weish1. Para essas instâncias, mais de uma combinação dos parâmetros  $(\alpha; \gamma; \epsilon)$  chegou ao resultado ótimo:

• weing1: 7 combinações.

• weing2: 15 combinações.

• weing3: 57 combinações.

• weish1: 5 combinações.

Em Anexo, são apresentadas as combinações dos parâmetros que proporcionam o ótimo em cada um desses problemas.

Em oito instâncias o AR não alcançou o ótimo com nenhuma das combinações de parâmetros analisadas. A Tabela 2 apresenta a melhor solução encontrada nesses casos e as respectivas combinações de parâmetros:

Os resultados apresentados na Tabela 2 apontam que o AR alcançou boas soluções ao se adotar determinadas configurações dos parâmetros  $\alpha$ ,  $\gamma$  e  $\epsilon$ . Nesse aspecto, torna-se relevante analisar a sensibilidade dos parâmetros do AR na tentativa de compreender como esses valores podem afetar a solução do problema da mochila [21].

Para realizar uma análise inicial, foram calculadas médias e desvios padrões a partir das melhores soluções encontradas por cada parâmetro ao longo das épocas. As Tabelas 3, 4 e 5, apresentam para os parâmetros  $\alpha$ ,  $\gamma$  e  $\epsilon$ , respectivamente, a menor e a maior média (média  $\pm$  desvio padrão) alcançada para cada um dos 12 problemas estudados.

Ao analisar a taxa de aprendizado,  $\alpha=0.01$  obteve a menor média em sete instâncias (weing1, weing2, weing3, weish1, weish2, weish30 e pb2) e  $\alpha=0.99$  conseguiu a maior média em cinco problemas (weing1, weing2, weing3, hp2 e pb1). Assim, uma hipótese é que a velocidade das atualizações do AR influencia diretamente da busca por boas soluções do PMM.

Já ao observar o fator de desconto,  $\gamma=0.99$  alcançou a maior média em quatro instâncias (sento1, hp1, hp2 e pb1). No entanto, para três casos o melhor fator desconto foi  $\gamma=0.01$  (weing1, weing2 e weing3). Como o

Tabela 2: Melhor solução para problemas em que o AR não alcançou o ótimo.

Problema	Ótimo	Melhor	$\alpha$	$\gamma$	$\epsilon$
		Solução			
sento1	7772	6911	0,99	0,75	0
weing7	1095445	1094917	0,01	0,45	0,05
weish2	4536	4483	0,30	0,45	0,10
weish30	11191	9387	0,75	0,60	0,05
hp1	3418	3319	0,60	0,01	0,10
hp2	3186	2903	0,01	0,15	0,05
pb1	3090	3004	0,15	0,75	0,10
			0,45	0,99	0,10
pb2	3186	2991	0,90	0,90	0

Fonte: elaborado pelos autores com base nos experimentos realizados.

Tabela 3: Médias e desvios padrões para a taxa de aprendizado  $(\alpha)$ .

Problema	Menor Média	土	Desvio Padrão	$\alpha$	Maior Média	$\pm$	Desvio Padrão	$\alpha$
sento1	5479	$\pm$	326	0,75	5598	土	313	0,30
weing1	131467	$\pm$	12143	0,01	138906	$\pm$	1628	0,99
weing2	118169	$\pm$	10893	0,01	128859	$\pm$	2310	0,99
weing3	83661	$\pm$	8493	0,01	93060	$\pm$	2457	0,99
weing7	1081612	$\pm$	7196	0,90	1084686	$\pm$	5527	0,15
weish1	4170	$\pm$	117	0,01	4259	$\pm$	117	0,30
weish2	4099	$\pm$	109	0,01	4164	$\pm$	114	0,45
weish30	8600	$\pm$	140	0,01	8718	$\pm$	206	0,75
hp1	3067	$\pm$	91	0,45	3098	$\pm$	96	0,01
hp2	2545	$\pm$	128	0,45	2572	$\pm$	120	0,99
pb1	2761	$\pm$	90	0,75	2776	$\pm$	97	0,99
pb2	2746	$\pm$	96	0,01	2805	$\pm$	58	0,90

Fonte: elaborado pelos autores com base nos experimentos realizados.

Tabela 4: Médias e desvios padrões para o fator de desconto  $(\gamma)$ .

Problema	Menor Média	$\pm$	Desvio Padrão	$\gamma$	Maior Média	$\pm$	Desvio Padrão	$\gamma$
sento1	5440	±	269	0,30	5643	±	262	0,99
weing1	131293	$\pm$	10549	0,99	139695	$\pm$	1520	0,01
weing2	117649	$\pm$	11618	0,99	129205	$\pm$	3454	0,01
weing3	83592	$\pm$	9525	0,99	92233	$\pm$	4592	0,01
weing7	1079628	$\pm$	7930	0,99	1084573	$\pm$	5741	0,60
weish1	4184	$\pm$	148	0,99	4281	$\pm$	109	0,75
weish2	4100	$\pm$	108	0,01	4208	$\pm$	114	0,75
weish30	8576	$\pm$	138	0,01	8811	$\pm$	238	0,90
hp1	3061	$\pm$	90	0,45	3142	$\pm$	88	0,99
hp2	2532	$\pm$	102,8	0,60	2621	$\pm$	135	0,99
pb1	2749	$\pm$	81,8	0,01	2835	$\pm$	88	0,99
pb2	2783	$\pm$	63	0,01	2791	$\pm$	89	0,45

Fonte: elaborado pelos autores com base nos experimentos realizados.

parâmetro  $\gamma$  controla a importância das recompensas futuras, uma hipótese é que a estrutura de reforços influenciou de forma distinta em cada instância.

Por último, os resultados para a política  $\epsilon-greedy$  revelam uma predominância de  $\epsilon=0,10$ , alcançando a maior média em 10 problemas (weing1, weing2, weing3, weish1, weish2, weish30, hp1, hp2, pb1 e pb2). Por outro lado,  $\epsilon=0$  obteve as piores médias em 11 instâncias (sento1, weing1, weing2, weing3, weish1, weish2, weish30, hp1, hp2, pb1 e pb2). Isso reforça a importância da adoção de uma política exploratória que adote uma

Tabela 5: Médias para a política  $\epsilon - greedy(\epsilon)$ .

Tubbia 5. Wedias para a pontiba e 9, 60ag (c).								
Problema	Menor Média	$\pm$	Desvio Padrão	$\epsilon$	Maior Média	$\pm$	Desvio Padrão	$\epsilon$
sento1	5507	±	335	0	5558	±	282	0,05
weing1	13633	$\pm$	10594	0	139036	$\pm$	1958	0,10
weing2	121671	$\pm$	12187	0	128660	$\pm$	2561	0,10
weing3	85473	$\pm$	9524	0	92312	$\pm$	3865	0,10
weing7	1082206	$\pm$	5729	0,10	1084175	$\pm$	5678	0,05
weish1	4163	$\pm$	126	0	4279	$\pm$	96	0,10
weish2	4104	$\pm$	131	0	4186	$\pm$	102	0,10
weish30	8611	$\pm$	188	0	8708	$\pm$	209	0,10
hp1	3028	$\pm$	89	0	3123	$\pm$	80	0,10
hp2	2456	$\pm$	69	0	2626	$\pm$	101	0,10
pb1	2720	$\pm$	97	0	2806	$\pm$	78	0,10
pb2	2729	$\pm$	96	0	2820	$\pm$	47	0,10

Fonte: elaborado pelos autores com base nos experimentos realizados.

parcela de ações aleatórias durante o aprendizado.

Assim, uma hipótese é que cada problema pode apresentar particularidades quanto a sensibilidade dos parâmetros. Nesse sentido, a modelagem matemática via Metodologia de Superfície de Resposta [21] é abordada na próxima seção na tentativa de verificar se existe uma correlação entre os parâmetros para as instâncias em que o AR não alcançou o ótimo.

## 6.2 Modelagem via metodologia de superfície de resposta

A Metodologia de Superfície de Resposta, em inglês *Response Surface Methodology* (RSM), reúne um conjunto de técnicas estatísticas para a otimização de processos [29]. A medida de desempenho é denominada resposta. Já as variáveis de entrada são ditas variáveis independentes (VIs) [29]. A Equação 7 apresenta a estrutura do modelo RSM de segunda ordem com duas VIs ( $x_1$  e  $x_2$ ):

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_1^2 + \beta_4 x_2^4 + \beta_5 x_1 x_2 + e, \tag{7}$$

em que, os coeficientes do modelo são representados por  $\beta_n$  e o efeito do erro (resíduo) na resposta é denotado por e.

O trabalho de [21] propõe a modelagem via RSM para análise dos efeitos dos parâmetros  $(\alpha, \gamma)$  do AR na solução do Problema do Caixeiro Viajante. Nesse sentido, foi proposto o modelo da Equação 8 [21]:

$$\hat{y} = \beta_0 + \beta_1 \alpha + \beta_2 \gamma + \beta_3 \alpha^2 + \beta_4 \gamma^2 + \beta_5 \alpha \gamma, \tag{8}$$

em que, os parâmetros  $\alpha$  e  $\gamma$  são as VIs da estrutura apresentada. Já  $\hat{y}$  é a variável resposta do modelo ajustado (medida de desempenho). O erro (e) não aparece na Equação 8, pois  $e_i$  é a diferença entre uma observação  $(y_i)$  e a respectiva resposta predita  $(\hat{y_i})$ , ou seja,  $e_i = y_i - \hat{y_i}$ .

Dessa forma, baseando-se em [21], neste trabalho é adotada a modelagem RSM para analisar a relação entre os parâmetros ( $\alpha$  e  $\gamma$ ) na solução das instâncias em que o AR não chegou ao resultado ótimo para o problema da mochila. Nesse aspecto, foram ajustados 8 modelos RSM (um para cada instância sem resultado ótimo) na estrutura da Equação 8. Para o ajuste dos modelos foram adotados os dados dos resultados de médias de solução das combinações de parâmetros ao longo de cada uma das épocas. Além disso, como o modelo da Equação 8 é estruturado apenas para  $\alpha$  e  $\gamma$ , foram adotados os dados relativos as combinações de parâmetros com  $\epsilon=0.01$ . Isso porque, conforme pode-se observar na Tabela 5 e evidenciado na seção anterior,  $\epsilon=0.01$  alcançou a maior média de solução em 10 problemas. Além disso, foi adotado o pacote RSM do *software* R [30, 31]. A Tabela 6 apresenta os coeficientes ajustados para cada um dos modelos.

Para avaliar a adequação dos modelos aos dados são analisadas algumas medidas de ajuste. A primeira medida de adequação busca verificar se os resíduos dos modelos seguem uma distribuição normal. Para isso,

Tabela 6: Coeficientes ajustados para os modelos RSM.

rabela of Coefficientes ajastados para os modelos resivi.							
Problema	$\beta_0$	$\beta_1$	$\beta_2$	$\beta_3$	$\beta_4$	$\beta_5$	
sento1	3291,091	339,578	35,033	-135,558	-50,004	-252,431	
weing7	997352	45466	108951	-23706	-105186	-68119	
weish2	3248,84	226,56	301,21	-113,90	-330,57	-419,27	
weish30	7246,00	436,67	1327,47	-201,31	-1260,25	-1104,07	
hp1	2564,64	-304,96	-65,28	213,70	105,62	-87,89	
hp2	1893,76	23,96	-188,02	-231,44	198,07	28,44	
pb1	2182,29	-60,98	149,24	140,57	-28,48	-346,26	
pb2	2285,37	-167,68	-31,05	82,77	65,77	-76,30	

Fonte: elaborado pelos autores com base nos experimentos realizados.

foi adotado o teste de normalidade de Kolmogorov-Smirnov (KS) [32]. Nesse teste, a hipótese inicial é que os resíduos seguem uma distribuição normal ( $p_{ks} > 0.05$ ), e a hipótese alternativa é que não seguem ( $p_{ks} < 0.05$ ). Outro teste é a análise de significância do modelo. Nesse teste, a hipótese inicial é que o modelo não é significante (p - valor > 0.05), e significante se aceita a hipótese alternativa (p - valor < 0.05). A Tabela 7 apresenta as medidas de ajustas analisadas para cada um dos modelos propostos.

Tabela 7: Medidas de adequação dos modelos RSM.

1 3					
Problema	p-valor	$p_{ks}$			
sento1	$< 2.2 \times 10^{-16}$	0,9688			
weing7	$< 2.2 \times 10^{-16}$	0,0756			
weish2	$< 2.2 \times 10^{-16}$	0,0002			
weish30	$< 2.2 \times 10^{-16}$	0,0029			
hp1	$< 2.2 \times 10^{-16}$	0,2171			
hp2	$< 2.2 \times 10^{-16}$	0,5654			
pb1	$< 2.2 \times 10^{-16}$	0,1087			
pb2	$7,406 \times 10^{-11}$	0,9101			

Fonte: elaborado pelos autores com base nos experimentos realizados.

As medidas adequação apresentadas na Tabela 7 revelam que todos os modelos são significantes, ou seja, p-valor<0.05. Quanto a suposição de normalidade ( $p_{ks}>0.05$ ) foi satisfeita para 6 modelos: sento1, weing7, hp1, hp2, pb1 e pb2. Assim, apenas para os modelos das instâncias weish2 e weish30, o teste KS indica que não existe normalidade dos resíduos.

Após verificar as medidas de adequação dos modelos foram realizadas análises adotando as ferramentas gráficas da RSM: gráfico de contornos e superfície de resposta. O gráfico de contornos apresenta em duas dimensões a relação entre as variáveis dos modelos: taxa de aprendizado (eixo X), fator de desconto (eixo Y) e a variável resposta (linhas de contornos). Já a superfície de resposta apresenta em três dimensões a relação entre as variáveis do modelo. Dessa forma, por esses gráficos é possível visualizar regiões em que a medida de desempenho tende a ser maximizada/minimizada pelas variáveis independentes ( $\alpha$  e  $\gamma$ ).

As Figuras 1 e 2 apresentam os gráficos de contornos para os modelos das instâncias sento1 e weing7, respectivamente. Por esses gráficos é possível notar como a medida de desempenho (linhas de contornos) comporta-se de maneira distinta nos dois modelos de acordo com os valores de  $\alpha$  e  $\gamma$ . Por exemplo, pelo modelo para a instância sento1 a resposta ajustada tende a se maximizar quando  $\alpha \to 1$  e  $\gamma \to 0$  (região rosa da Figura 1). Já para o problema weing7, a resposta tende a ser maximizada em uma região de parâmetros centralizada no gráfico de contornos (região rosa da Figura 2).

As Figuras 3, 4, 5 e 6 apresentam, por sua vez, as superfícies de resposta para os modelos referentes as instâncias hp1, hp2, pb1 e pb2, respectivamente. Por esses gráficos é possível observar como em geral os parâmetros contribuiram de maneira distinta na solução dessas instâncias. No entanto, pode-se observar para esses 4 modelos uma tendência em maximizar a resposta ajustada para  $\alpha \to 0$  e  $\gamma \to 1$ . Além disso, vale ressaltar a semelhança entre as superfícies dos modelos das instâncias hp1 (Figura 3) e pb1 (Figura 6).

Figura 1: Gráfico de contornos (2D) para o modelo RSM referente ao problema sento1.

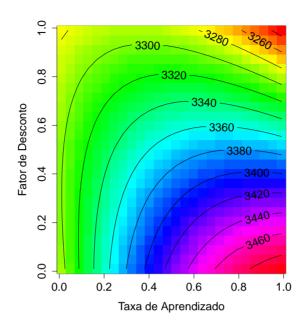
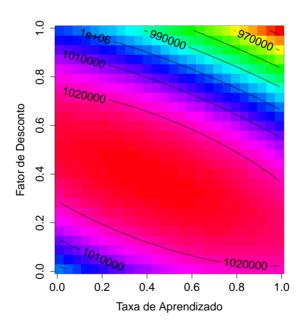


Figura 2: Gráfico de contornos (2D) para o modelo RSM referente ao problema weing7.



Assim, a modelagem matemática via RSM reforça a hipótese de que cada instância do problema da mochila pode apresentar particularidades quanto a sensibilidade dos parâmetros do AR. No entanto, a partir das ferramentas gráficas da RSM é possível identificar regiões de parâmetros que tendem a maximizar a resposta ajustada para modelos de instâncias distâncias, conforme Figuras 3, 4, 5 e 6.

Figura 3: Superfície de resposta (3D) para o modelo RSM referente ao problema hp1.

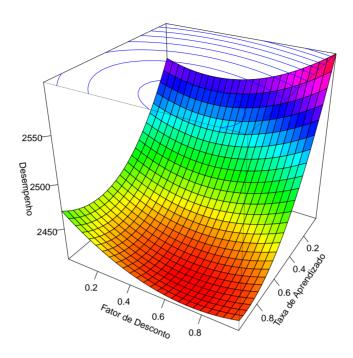


Figura 4: Superfície de resposta (3D) para o modelo RSM referente ao problema hp2.

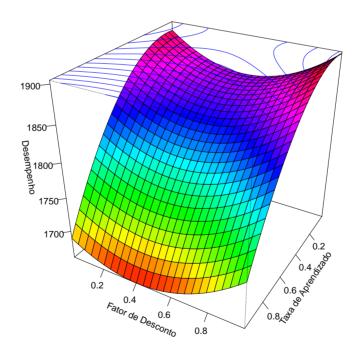


Figura 5: Superfície de resposta (3D) para o modelo RSM referente ao problema pb1.

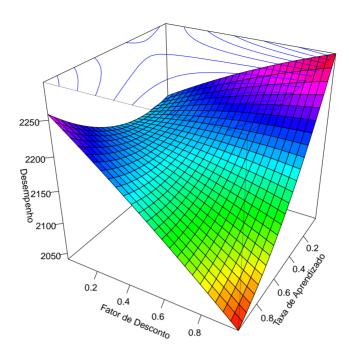
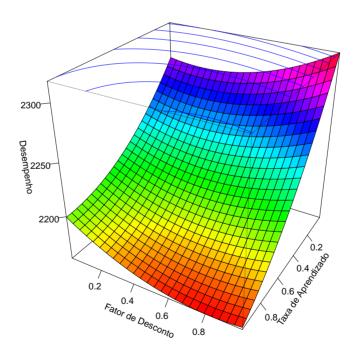


Figura 6: Superfície de resposta (3D) para o modelo RSM referente ao problema pb2.



#### 7 Conclusão

Neste trabalho, o objetivo foi analisar o desempenho do Aprendizado por Reforço na Solução do Problema da Mochila Multidimensional. Para isso, foi proposto um modelo de AR em estados, ações e recompensas. O algoritmo Q-learning alcançou a solução ótima em 4 instâncias: weing1, weing2, weing3 e weish1. Desses quatro problemas, weish1 é o que possui o maior espaço de busca, composto por  $5 \times 2^{30}$  soluções possíveis. Já para as outras 8 instâncias, o Q-learning não alcançou a solução ótima. Nesse sentido, foram realizadas análises de sensibilidade dos parâmetros do AR na tentativa de compreender como  $\alpha$ ,  $\gamma$  e  $\epsilon$  influenciam na resolução do Problema da Mochila Multidimensional.

A modelagem matemática via modelos superfície de resposta aponta que as instâncias do PMM podem apresentar particularidades quanto aos valores adotados dos parâmetros do AR. Nesse sentido, a partir dos gráficos da RSM é possível visualizar para cada instância o comportamento da resposta (medida de desempenho) de acordo com  $\alpha$  e  $\gamma$ . Nessa linha, é possível encontrar uma mesma região que maximize a resposta ajustada para instâncias distintas, conforme evidenciando para os problemas hp1, hp2, pb1 e pb2. Dessa forma, a RSM é uma ferramenta que auxilia na visualização de regiões de parâmetros do AR que tendem a ter melhores soluções nas instâncias do PMM.

Em trabalhos futuros, serão aplicados outros métodos de AR na resolução do PMM, como o SARSA. Também serão comparados o desempenho da solução do PMM entre métodos de AR com outros algoritmos tradicionais, como Programação Dinâmica, Algoritmos Genéticos e Busca Tabu. Nesse sentido, espera-se analisar o real esforço computacional entre os algoritmos de AR e esses outros métodos de comum aplicação na resolução do Problema da Mochila. Além disso, o comportamento dos parâmetros do AR será investigado em outros problemas de otimização combinatória.

# Agradecimentos

Agradecemos à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), Fundação de Amparo a Pesquisa do Estado de Minas Gerais (FAPEMIG), Universidade Federal de São João del-Rei (UFSJ), Centro Federal de Educação Tecnológica de Minas Gerais (CEFET-MG) e Programa de Pós-Graduação em Engenharia Elétrica (PPGEL - UFSJ/CEFET-MG) pelo apoio.

### Referências

- [1] CHU, P.; BEASLEY, J. A genetic algorithm for the multidimensional knapsack problem. *Journal of Heuristics*, v. 4, n. 1, p. 63–86, 1998. ISSN 1572-9397. Disponível em: <a href="http://dx.doi.org/10.1023/A:1009642405419">http://dx.doi.org/10.1023/A:1009642405419</a>.
- [2] FRÉVILLE, A. The multidimensional 0–1 knapsack problem: An overview. *European Journal of Operational Research*, v. 155, n. 1, p. 1 21, 2004. ISSN 0377-2217. Disponível em: <a href="http://dx.doi.org/10.1016/S0377-2217(03)00274-1">http://dx.doi.org/10.1016/S0377-2217(03)00274-1</a>.
- [3] HANAFI, S.; FREVILLE, A. An efficient tabu search approach for the 0–1 multidimensional knapsack problem. *European Journal of Operational Research*, v. 106, n. 2, p. 659 675, 1998. ISSN 0377-2217. Disponível em: <a href="http://dx.doi.org/10.1016/S0377-2217(97)00296-8">http://dx.doi.org/10.1016/S0377-2217(97)00296-8</a>.
- [4] KRAUSE, J.; CORDEIRO, J. A.; LOPES, H. S. Comparação de métodos de computação evolucionária para o problema da mochila multidimensional. In: LOPES, H. S.; RODRIGUES, L. C. de A.; STEINER, M. T. A. (Ed.). *Meta-Heurísticas em Pesquisa Operacional*. 1. ed. Curitiba, PR: Omnipax, 2013. cap. 6, p. 87–98. ISBN 978-85-64619-10-4. Disponível em: <a href="http://dx.doi.org/10.7436/2013.mhpo.06">http://dx.doi.org/10.7436/2013.mhpo.06</a>.
- [5] MARQUES, F. d. P.; ARENALES, M. N. O Problema da mochila compartimentada e aplicações. *Pesquisa Operacional*, scielo, v. 22, p. 285 304, 07 2002. ISSN 0101-7438. Disponível em: <a href="http://www.scielo.br/scielo.php?script=sci\_arttext&pid=S0101-74382002000300001&nrm=iso>">http://www.scielo.br/scielo.php?script=sci\_arttext&pid=S0101-74382002000300001&nrm=iso>">http://www.scielo.br/scielo.php?script=sci\_arttext&pid=S0101-74382002000300001&nrm=iso>">http://www.scielo.br/scielo.php?script=sci\_arttext&pid=S0101-74382002000300001&nrm=iso>">http://www.scielo.br/scielo.php?script=sci\_arttext&pid=S0101-74382002000300001&nrm=iso>">http://www.scielo.br/scielo.php?script=sci\_arttext&pid=S0101-74382002000300001&nrm=iso>">http://www.scielo.br/scielo.php?script=sci\_arttext&pid=S0101-74382002000300001&nrm=iso>">http://www.scielo.br/scielo.php?script=sci\_arttext&pid=S0101-74382002000300001&nrm=iso>">http://www.scielo.br/scielo.php?script=sci\_arttext&pid=S0101-74382002000300001&nrm=iso>">http://www.scielo.php?script=sci\_arttext&pid=S0101-74382002000300001&nrm=iso>">http://www.scielo.php?script=sci\_arttext&pid=S0101-74382002000300001&nrm=iso>">http://www.scielo.php?script=sci\_arttext&pid=S0101-74382002000300001&nrm=iso>">http://www.scielo.php?script=sci\_arttext&pid=S0101-74382002000300001&nrm=iso>">http://www.scielo.php?script=sci\_arttext&pid=S0101-74382002000300001&nrm=iso>">http://www.scielo.php?script=sci\_arttext&pid=S0101-74382002000300001&nrm=iso>">http://www.scielo.php?script=sci\_arttext&pid=S0101-74382002000300001&nrm=iso>">http://www.scielo.php?script=sci\_arttext&pid=S0101-74382002000300001&nrm=iso>">http://www.scielo.php?script=sci\_arttext&pid=S0101-74382002000300001&nrm=iso>">http://www.scielo.php?script=sci\_arttext&pid=S0101-74382002000300001&nrm=iso>">http://www.scielo.php?script=sci\_arttext&pid=S0101-74382002000300001&nrm=iso>">http://www.scielo.php?script=sci\_arttext&pid=S0101-74382002000300001&nrm=iso>">http://www.scielo.php?script=sci\_arttext&pid=S0101-743820000001&nrm=iso>">http://www.scielo.php?script=s

- [6] JUNIOR, J. A. L. S.; SILVA, D. J. A.; OLIVEIRA, R. C. L. Parametrização de um algoritmo cultural híbrido para o problema da mochila multidimensional. *XII SBAI Simpósio Brasileiro de Automação Inteligente*, p. 678–688, 2015. Disponível em: <a href="http://www.sbai2015.dca.ufrn.br/download/artigo/193">http://www.sbai2015.dca.ufrn.br/download/artigo/193</a>.
- [7] LIU, W.; LIU, G. Genetic algorithm with directional mutation based on greedy strategy for large-scale 0-1 knapsack problems. *International Journal of Advancements in Computing Technology*, v. 4, n. 3, p. 66–74, 2012. Disponível em: <a href="http://dx.doi.org/10.4156/ijact.vol4.issue3.9">http://dx.doi.org/10.4156/ijact.vol4.issue3.9</a>>.
- [8] LU, T.-C.; YU, G.-R. An adaptive population multi-objective quantum-inspired evolutionary algorithm for multi-objective 0/1 knapsack problems. *Information Sciences*, v. 243, p. 39–56, 2013. Disponível em: <a href="http://dx.doi.org/10.1016/j.ins.2013.04.018">http://dx.doi.org/10.1016/j.ins.2013.04.018</a>>.
- [9] RONG, A.; FIGUEIRA, J.; KLAMROTH, K. Dynamic programming based algorithms for the discounted 0-1 knapsack problem. *Applied Mathematics and Computation*, v. 218, n. 12, p. 6921–6933, 2012. Disponível em: <a href="http://dx.doi.org/10.1016/j.amc.2011.12.068">http://dx.doi.org/10.1016/j.amc.2011.12.068</a>>.
- [10] YANG, Z.; WANG, G.; CHU, F. An effective grasp and tabu search for the 0-1 quadratic knapsack problem. *Computers and Operations Research*, v. 40, n. 5, p. 1176–1185, 2013. Disponível em: <a href="http://dx.doi.org/10.1016/j.cor.2012.11.023">http://dx.doi.org/10.1016/j.cor.2012.11.023</a>.
- [11] KONG, M.; TIAN, P.; KAO, Y. A new ant colony optimization algorithm for the multidimensional knapsack problem. *Computers & Operations Research*, v. 35, n. 8, p. 2672 2683, 2008. ISSN 0305-0548. Queues in Practice. Disponível em: <a href="http://dx.doi.org/10.1016/j.cor.2006.12.029">http://dx.doi.org/10.1016/j.cor.2006.12.029</a>>.
- [12] KE, L. et al. An ant colony optimization approach for the multidimensional knapsack problem. *Journal of Heuristics*, v. 16, n. 1, p. 65–83, 2010. ISSN 1572-9397. Disponível em: <a href="http://dx.doi.org/10.1007/s10732-008-9087-x">http://dx.doi.org/10.1007/s10732-008-9087-x</a>.
- [13] CHEN, W.-N. et al. A novel set-based particle swarm optimization method for discrete optimization problems. *IEEE Transactions on Evolutionary Computation*, v. 14, n. 2, p. 278–300, 2010. Disponível em: <a href="http://dx.doi.org/10.1109/TEVC.2009.2030331">http://dx.doi.org/10.1109/TEVC.2009.2030331</a>>.
- [14] CHIH, M. et al. Particle swarm optimization with time-varying acceleration coefficients for the multidimensional knapsack problem. *Applied Mathematical Modelling*, v. 38, n. 4, p. 1338–1350, 2014. Disponível em: <a href="http://dx.doi.org/10.1016/j.apm.2013.08.009">http://dx.doi.org/10.1016/j.apm.2013.08.009</a>.
- [15] GLIMSDAL, S.; GRANMO, O.-C. Gaussian process based optimistic knapsack sampling with applications to stochastic resource allocation. In: *Proceedings of the 24th Midwest Artificial Intelligence and Cognitive Science Conference, MAICS 2013.* [s.n.], 2013. p. 43–50. Disponível em: <a href="https://www.scopus.com/inward/record.uri?eid=2-s2.0-84880565427&partnerID=40&md5=a68fbb8af893f5d8f0633896bb23ae34>.">https://www.scopus.com/inward/record.uri?eid=2-s2.0-84880565427&partnerID=40&md5=a68fbb8af893f5d8f0633896bb23ae34>.</a>
- [16] ARIN, A.; RABADI, G. Integrating estimation of distribution algorithms versus q-learning into meta-raps for solving the 0-1 multidimensional knapsack problem. *Computers & Industrial Engineering*, p. –, 2016. ISSN 0360-8352. Disponível em: <a href="http://dx.doi.org/10.1016/j.cie.2016.10.022">http://dx.doi.org/10.1016/j.cie.2016.10.022</a>.
- [17] SUTTON, R.; BARTO, A. Reinforcement Learning: An Introduction. 1st. ed. [S.l.]: Cambridge, MA: MIT Press, 1998.
- [18] RUSSELL, S. J.; NORVING, P. Artificial Intelligence. [S.l.]: Campus, 3st ed., 2013.
- [19] GAMBARDELLA, L. M.; DORIGO, M. Ant-Q: A reinforcement learning approach to the traveling salesman problem. *Proceedings of the 12th International Conference on Machine Learning*, p. 252–260, 1995. Disponível em: <a href="http://faculty.washington.edu/paymana/swarm/gambardella95-icml.pdf">http://faculty.washington.edu/paymana/swarm/gambardella95-icml.pdf</a>>.
- [20] LIU, F.; ZENG, G. Study of genetic algorithm with reinforcement learning to solve the TSP. *Expert Systems with Applications*, v. 36, n. 3, p. 6995 7001, 2009. Disponível em: <a href="http://dx.doi.org/10.1016/j.eswa.2008.08.026">http://dx.doi.org/10.1016/j.eswa.2008.08.026</a>.

- [21] OTTONI, A. L. C.; NEPOMUCENO, E. G.; OLIVEIRA, M. S. Análise de Sensibilidade dos Parâmetros do Aprendizado por Reforço na Solução do Problema do Caixeiro Viajante: Modelagem via Superfície de Resposta. *XXI CBA Congresso Brasileiro de Automática 2016*, 2016. Disponível em: <a href="http://www.swge.inf.br/PDF/CBA2016-0170\_047213.PDF">http://www.swge.inf.br/PDF/CBA2016-0170\_047213.PDF</a>.
- [22] OTTONI, A. L. C. et al. Análise do Desempenho do Aprendizado por Reforço na Solução do Problema do Caixeiro Viajante. *XII SBAI Simpósio Brasileiro de Automação Inteligente*, p. 43–48, 2015. Disponível em: <a href="http://www.sbai2015.dca.ufrn.br/download/artigo/17">http://www.sbai2015.dca.ufrn.br/download/artigo/17</a>>.
- [23] SCHWEIGHOFER, N.; DOYA, K. Meta-learning in reinforcement learning. *Neural Networks*, v. 16, n. 1, p. 5–9, 2003. Disponível em: <a href="http://dx.doi.org/10.1016/S0893-6080(02)00228-9">http://dx.doi.org/10.1016/S0893-6080(02)00228-9</a>.
- [24] EVEN-DAR, E.; MANSOUR, Y. Learning Rates for Q-learning. *Journal of Machine Learning Research*, v. 5, p. 1–25, 2003. Disponível em: <a href="http://www.jmlr.org/papers/volume5/evendar03a/evendar03a.pdf">http://www.jmlr.org/papers/volume5/evendar03a/evendar03a.pdf</a>>.
- [25] OTTONI, A. L. C. et al. Análise da influência da taxa de aprendizado e do fator de desconto sobre o desempenho dos algoritmos Q-learning e SARSA: aplicação do aprendizado por reforço na navegação autônoma. *Revista Brasileira de Computação Aplicada*, v. 8, n. 2, p. 44–59, 2016. Disponível em: <a href="http://dx.doi.org/10.5335/rbca.v8i2.5249">http://dx.doi.org/10.5335/rbca.v8i2.5249</a>>.
- [26] PELLEGRINI, J.; WAINER, J. Processos de Decisão de Markov: um tutorial. *RITA Revista de Informática Teórica e Aplicada*, v. 14, n. 2, p. 133–179, 2007. Disponível em: <a href="http://www.seer.ufrgs.br/rita/article/view/rita\_v14\_n2\_p133-179">http://www.seer.ufrgs.br/rita/article/view/rita\_v14\_n2\_p133-179</a>.
- [27] WATKINS, C. J.; DAYAN, P. Technical note Q-learning. *Machine Learning*, v. 8, n. 3, p. 279–292, 1992. Disponível em: <a href="http://dx.doi.org/10.1023/A:1022676722315">http://dx.doi.org/10.1023/A:1022676722315</a>>.
- [28] BEASLEY, J. E. Or-library: Distributing test problems by electronic mail. *Journal of the Operational Research Society*, v. 41, n. 11, p. 1069–1072, 1990. ISSN 1476-9360. Disponível em: <a href="http://dx.doi.org/10.1057/jors.1990.166">http://dx.doi.org/10.1057/jors.1990.166</a>>.
- [29] MYERS, R. H.; MONTGOMERY, D. C.; Anderson-Cook, C. M. Response surface methodology: process and product optimization using designed experiments. [S.l.]: John Wiley & Sons, 3 ed, 2009.
- [30] LENTH, R. V. Response-Surface Methods in R, using rsm. *Journal of Statistical Software*, v. 32, n. 7, p. 1–17, 2009. Disponível em: <a href="http://dx.doi.org/10.18637/jss.v032.i07">http://dx.doi.org/10.18637/jss.v032.i07</a>>.
- [31] R Core Team. *R: A Language and Environment for Statistical Computing*. Vienna, Austria, 2013. Disponível em: <a href="https://www.r-project.org/">https://www.r-project.org/</a>>.
- [32] LOPES, R. H. C. Kolmogorov-smirnov test. In: \_\_\_\_\_. *International Encyclopedia of Statistical Science*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2011. p. 718–720. ISBN 978-3-642-04898-2. Disponível em: <a href="http://dx.doi.org/10.1007/978-3-642-04898-2\_326">http://dx.doi.org/10.1007/978-3-642-04898-2\_326</a>.

#### Anexo A

Combinações dos parâmetros que proporcionam o ótimo são apresentadas na forma  $(\alpha; \gamma; \epsilon)$ :

- weing1: (0,01; 0,30; 0,05), (0,15; 0,45; 0,10), (0,30; 0,15; 0,10), (0,30; 0,30; 0,10), (0,45; 0,45; 0); (0,75; 0,15; 0,05); (0,90; 0,01; 0,05);
- weing2: (0,15; 0,30; 0,05), (0,30; 0,01; 0,10), (0,30; 0,30; 0,05), (0,30; 0,45; 0,05), (0,45; 0,01; 0,10), (0,45; 0,15; 0,05); (0,45; 0,15; 0,10); (0,45; 0,45; 0,05); (0,45; 0,30; 0,10), (0,60; 0,30; 0,10), (0,75; 0,15; 0,10), (0,75; 0,45; 0,10), (0,90; 0,75; 0,05), (0,99; 0,01; 0,10), (0,99; 0,30; 0,05);
- weing3: (0,01; 0,15; 0,10), (0,01; 0,90; 0,05), (0,15; 0,01; 0), (0,15; 0,01; 0,10), (0,15; 0,15; 0,05), (0,01; 0,30; 0,10), (0,01; 0,60; 0,10), (0,30; 0,01; 0,05), (0,30; 0,01; 0,10), (0,30; 0,30; 0,05), (0,30; 0,45; 0,10), (0,30; 0,60; 0,05), (0,30; 0,60; 0,10), (0,30; 0,75; 0,05), (0,45; 0,01; 0,05), (0,45; 0,01; 0,10), (0,45; 0,30; 0), (0,45; 0,30; 0,10), (0,45; 0,45; 0,05), (0,45; 0,45; 0,10), (0,45; 0,60; 0,10), (0,45; 0,90; 0,10), (0,60; 0,01; 0), (0,60; 0,01; 0,05), (0,60; 0,01; 0,10), (0,60; 0,30; 0,05), (0,60; 0,30; 0,10), (0,60; 0,30; 0,10), (0,75; 0,45; 0,10), (0,75; 0,45; 0,10), (0,75; 0,45; 0,10), (0,75; 0,45; 0,10), (0,75; 0,45; 0,10), (0,75; 0,45; 0,10), (0,75; 0,45; 0,10), (0,75; 0,45; 0,10), (0,90; 0,30; 0,10), (0,90; 0,45; 0,10), (0,90; 0,60; 0,05), (0,90; 0,75; 0), (0,90; 0,75; 0,05), (0,90; 0,90; 0,05), (0,90; 0,99; 0,10), (0,99; 0,01; 0), (0,99; 0,45; 0,10), (0,99; 0,75; 0,05), (0,99; 0,45; 0,10), (0,99; 0,45; 0,10), (0,99; 0,75; 0,05), (0,99; 0,99; 0);
- weish1: (0.15; 0.75; 0.10), (0.30; 0.99; 0.10), (0.45; 0.75; 0.10), (0.75; 0.60; 0.10), (0.90; 0.15; 0.10).