

# Comparação da Eficiência do Perceptron de Múltiplas Camadas e da Rede de Função de Base Radial como Modelos Substitutos na Otimização Multiobjetivo com Algoritmo Genético

Alexandre N. Barbosa<sup>1</sup>, Lamartine N. F. Guimarães<sup>2</sup>

<sup>1</sup>Instituto de Aeronáutica e Espaço – IAE  
Departamento de Ciência e Tecnologia Aeroespacial - DCTA

<sup>2</sup>Instituto de Estudos Avançados – IEAv  
Departamento de Ciência e Tecnologia Aeroespacial - DCTA

nogueiraanb@iae.cta.br, guimarae@ieav.cta.br

**Abstract.** *The surrogate models are functions with low computational cost used for replacing objective functions with high computational cost in order to increase the efficiency of the optimization reducing the number of evaluations of the objective function. This article suggests that the radial basis function network as surrogate model is more efficient than the multilayer perceptron when the objective function has many local maxima and minima.*

**Resumo.** *Os modelos substitutos são funções com baixo custo computacional usados para substituir funções objetivo com alto custo computacional, a fim de aumentar a eficiência da otimização, reduzindo o número de avaliações da função objetivo. Este artigo sugere que a rede de função radial base como modelo substituto é mais eficiente do que o perceptron de múltiplas camadas quando a função objetivo tem muitos máximos e mínimos locais.*

**Palavras-chave:** *Redes Neurais, Otimização Multiobjetivo e Algoritmo Genético.*

## 1. Introdução

Os modelos substitutos são funções com custo computacional baixo, usados para substituir funções objetivo com alto custo computacional a fim de aumentar a eficiência da otimização. No entanto, a escolha do modelo mais adequado requer uma análise de sua eficiência. Se nenhuma análise for feita, o uso do modelo poderá ser inexpressivo. Portanto, o objetivo deste trabalho é comparar a eficiência de duas redes neurais como modelos substitutos, a fim de estudar um procedimento que permita a escolha da rede neural mais adequada. Para tanto, experimentam-se as redes neurais na otimização multiobjetivo de duas funções algébricas, obtidas na literatura. Comparam-se o perceptron de múltiplas camadas (MLP) e a rede de função de base radial (RBF). Dentre as redes neurais, estas são as mais empregadas como modelos substitutos [Naidu 2004]. Os resultados sugerem que a segunda é mais eficiente do que a primeira quando a função objetivo tem muitos máximos e mínimos locais.

## 2. Metodologia

Utiliza-se o algoritmo genético não geracional multiobjetivo conjuntamente com as redes neurais. A particularidade deste algoritmo, a qual é levada em conta na análise comparativa da eficiência dos modelos substitutos, é a de gerar somente dois indivíduos “filhos” a cada geração a partir de dois indivíduos “pais” [Borges 1999]. Além disso, considera-se, também, o tamanho da população e a quantidade de gerações, os quais são os principais parâmetros de ajuste do algoritmo genético.

Compara-se a eficiência das redes neurais em relação à otimização sem elas. A rede neural é eficiente quando, sem ela, não se é capaz de gerar uma quantidade de soluções Pareto-ótimo igual ou superior à gerada com o seu auxílio sem ter que exceder em quantidade de ativações da função objetivo. Quanto maior for essa necessidade, maior será a eficiência da rede neural.

O ganho de eficiência é quantificado da seguinte forma:

$$\eta = 2 \cdot (q_2 - q_1) \quad (1)$$

$\eta$ : acréscimo de ativações na otimização sem a rede neural para ter-se uma quantidade de soluções Pareto-ótimo igual ou superior à obtida com o seu auxílio

$q_1$ : máximo de gerações, sem o auxílio da rede neural, tal que a quantidade de ativações da função objetivo seja menor ou igual à de ativações com o seu auxílio

$q_2$ : mínimo de gerações, sem o auxílio da rede neural, tal que a quantidade de soluções Pareto-ótimo seja maior ou igual à de soluções obtidas com o seu auxílio

A quantidade  $q_1$  é determinada em função da particularidade e dos parâmetros de ajuste do algoritmo genético, já mencionados, e também em função do algoritmo de otimização com modelos substitutos, descrito em quatro passos. Primeiro, avalia-se a função objetivo, certa quantidade de vezes, a fim de gerar exemplos suficientes para construir o modelo substituto. Segundo, realiza-se a otimização por ciclos, definindo uma probabilidade de substituição da função objetivo pelo modelo substituto, constante ou maior a cada ciclo. Terceiro, no final de cada ciclo, inserem-se as soluções avaliadas com a função objetivo no conjunto de padrões de treinamento e constroi-se, novamente, o modelo substituto. Quarto, após o último ciclo de otimização, avalia-se com a função objetivo as soluções Pareto-ótimo encontradas com o auxílio do modelo substituto. Assim, a quantidade  $q_1$  é expressa pela Equação 2.

$$q_1 = \frac{(1-P) \cdot N}{2} \cdot p + (1-P) \cdot g + \frac{1}{2} \cdot e \quad (2)$$

P: valor médio da probabilidade de substituição

N: quantidade de ciclos de otimização

p: tamanho da população

g: quantidade de gerações

e: quantidade de exemplos iniciais

Determina-se a quantidade  $q_2$ , primeiro, atribuindo-se um valor inicial a ela. Em seguida, como segundo passo, varia-se  $q_2$  em incrementos de valor constante. Faz-se, então, a otimização sem o auxílio das redes neurais, considerando a quantidade de gerações igual à  $q_2$ . Retorna-se ao segundo passo enquanto a quantidade de soluções Pareto-ótimo for menor do que a quantidade obtida com o auxílio da rede neural.

### 3. Estudo de Caso

O estudo de caso consistiu em minimizar as funções algébricas descritas na Tabela 1 com o auxílio da MLP e RBF, para, em seguida, comparar a eficiência destas.

**Tabela 1. Descrição das funções algébricas.**

Função	Componentes	Restrições
F1 (*)	$f_1 = (x_1 - 2)^2 + (x_2 - 1)^2 + 2$ $f_2 = 9 \cdot x_1 - (x_2 - 1)^2$	$-20 \leq x_i \leq 20, i = 1, 2$
F2 (**)	$f_1(x_1, x_2) = 1 + (A_1 - B_1)^2 + (A_2 - B_2)^2$ $f_2(x_1, x_2) = (x_1 + 3)^2 + (x_2 + 1)^2$ $A_1 = 0.5 \cdot \sin 1 - 2 \cdot \cos 1 + \sin 2 - 1.5 \cdot \cos 2$ $A_2 = 1.5 \cdot \sin 1 - \cos 1 + 2 \cdot \sin 2 - 0.5 \cdot \cos 2$ $B_1 = 0.5 \cdot \sin x_1 - 2 \cdot \cos x_1 + \sin x_2 - 1.5 \cdot \cos x_2$ $B_2 = 1.5 \cdot \sin x_1 - \cos x_1 + 2 \cdot \sin x_2 - 0.5 \cdot \cos x_2$	$-\pi \leq x_i \leq \pi, i = 1, 2$

(\*) [Galski 2006]/ (\*\*) [Deb et al. 2000]

Ajustaram-se os parâmetros da MLP com os mesmos valores para a aproximação de ambas as funções. Como função de ativação, usou-se a logística com coeficiente igual a 2. A arquitetura foi definida com uma camada escondida contendo 10 neurônios. A taxa de aprendizagem foi de 0.1. O momento foi de 0.5. A quantidade máxima de épocas de treinamento foi de 3000. E, a variação mínima do erro quadrático médio foi de 0.0001.

O ajuste da RBF foi diferente para cada função. Apenas a função de base radial foi a mesma, a Gaussiana. Para a aproximação da F1, o desvio padrão adotado foi de 25 e a quantidade de centros foi de 70. Enquanto, para a aproximação de F2, o desvio padrão foi de 2.5 e a quantidade de centros foi de 40.

Os parâmetros da otimização, adotados em cada caso, estão na Tabela 2.

**Tabela 2. Parâmetros da otimização.**

Parâmetro	Valor			
	F1		F2	
	MLP	RBF	MLP	RBF
Tamanho da população (p)	70	70	70	70
Quantidade de gerações (g)	150	150	450	300
Método de amostragem	LHS	LHS	LHS	LHS
Quantidade de exemplos iniciais (e)	15	80	30	80
Ciclos de otimização (N)	3	3	3	3
Probabilidade de substituição no 1º ciclo	0.94	0.94	0.90	0.90
Probabilidade de substituição no 2º ciclo	0.96	0.96	0.93	0.93
Probabilidade de substituição no 3º ciclo	1	1	1	1
Probabilidade de substituição média (P)	0.967	0.967	0.944	0.944

LHS: *Latin Hypercube Sampling*

Após a realização de 10 ensaios, obteve-se a quantidade média de soluções Pareto-ótimo para cada caso (Tabela 3).

**Tabela 3. Quantidade média de soluções Pareto-ótimo.**

	MLP	RBF
F1	52	52
F2	12	39

#### 4. Resultados

Esta análise comparativa sugere que a RBF seja mais eficiente do que a MLP quando a função objetivo tem muitos máximos e mínimos locais. A Função F2 apresenta uma quantidade de máximos e mínimos locais maior do que a da F1. No caso da F2, a necessidade de exceder em quantidade de ativações na otimização sem o auxílio da rede neural foi maior para alcançar-se os resultados da RBF, sugerindo que, neste caso, esta seja mais eficiente do que a MLP. Por outro lado, no caso da F1, a RBF não se mostrou mais eficiente do que a MLP (Tabela 4).

**Tabela 4. Ganho de eficiência.**

	MLP	RBF
F1	128	60
F2	116	230

#### 5. Conclusões

O objetivo deste trabalho foi comparar a eficiência de redes neurais como modelos substitutos em problemas de otimização multiobjetivo. Para tanto, experimentaram-se os modelos substitutos na otimização multiobjetivo de duas funções algébricas, obtidas na literatura. Compararam-se duas redes: o perceptron de múltiplas camadas (MLP) e a rede de função de base radial (RBF). Os resultados mostram que a rede mais apropriada depende de uma análise da função objetivo. Sugere-se que a RBF seja mais eficiente do que a MLP quando as componentes da função multiobjetivo possuem uma quantidade maior de máximos e mínimos locais.

#### Referências

- Borges, C. C. H. (1999) Algoritmos Genéticos para Otimização em Dinâmica de Estruturas. 216 p. Universidade Federal do Rio de Janeiro (UFRJ).
- Deb, K.; Pratap, A.; Agarwal S.; Meyarivan T. (2000) A Fast and Elitist Multi-Objective Genetic Algorithm: NSGA-II. Indian Institute of Technology, Kanpur.
- Galski, R. L. (2006) Desenvolvimento de Versões Aprimoradas, Híbridas, Paralela e Multiobjetivo do Método da Otimização Extrema Generalizada e sua Aplicação no Projeto de Sistemas Espaciais. 282 p. Instituto Nacional de Pesquisas Espaciais (INPE).
- Naidu, S. L. S. (2004) Neural Network Surrogate Model For Multidisciplinary Design Optimization. 100 p. Dissertation (Master of Technology) – India Institute of Technology, Bombay.