

Uso do Perceptron como Modelo Substituto na Otimização Multiobjetivo com Algoritmo Genético

Alexandre N. Barbosa¹, José Demisio S. Silva², Rafael D. C. dos Santos²

¹Instituto de Aeronáutica e Espaço (IAE)
Departamento de Ciência e Tecnologia Aeroespacial (DCTA), São J. dos Campos, SP

²Laboratório Associado de Computação e Matemática Aplicada (LAC)
Instituto Nacional de Pesquisas Espaciais (INPE), São José dos Campos, SP
nogueiraanb@iae.cta.br, {demisio, rafael}.santos@lac.inpe.br

Resumo. *A combinação de técnicas de inteligência artificial pode produzir resultados surpreendentes. Neste trabalho, uma rede perceptron é usada para guiar o processo de otimização, realizado por um algoritmo genético, assumindo o papel da função objetivo original, uma quantidade de vezes, a fim de exigir menos avaliações desta função. Este recurso se aplica em problemas onde: a função objetivo original tem alto custo computacional; o gradiente da função não é conhecido; e, a otimização não é viável no intervalo de tempo em que se deseja executá-la, mesmo considerando os recursos computacionais e a aplicação da computação paralela. Em tal aplicação, a técnica de otimização com modelos substitutos tem desempenho melhor do que a otimização sem a utilização deste recurso e, além disso, isto ocorre sem interferir na alta fidelidade dos resultados finais da otimização.*

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

1. Otimização com Modelo Substituto

Em geral, uma das dificuldades da otimização de um sistema complexo é a de ter que lidar com uma função objetivo com custo computacional elevado, cujo gradiente não é facilmente dedutível, pois a função é avaliada por rotinas computacionais complexas que simulam, com alta fidelidade, o comportamento do sistema. A abordagem convencional para superar esta dificuldade envolve o uso da computação paralela. Todavia, a computação paralela requer a utilização de meios computacionais que encarecem o seu emprego. Porém, há formas de se obter bons resultados, combinadas com a abordagem convencional, que implicam um gasto menor com meios computacionais e que, além disso, melhoram a eficiência. Por exemplo, o uso de modelos com custo computacional baixo que possam substituir, parcialmente ou totalmente, a função objetivo original durante o processo de otimização. Na literatura, se define uma probabilidade de substituição, a fim de se manter certo compromisso com a alta fidelidade do modelo original, durante o processo de otimização, uma vez que o modelo substituto é uma aproximação do original.

A abordagem deste trabalho consiste em fazer com que uma rede perceptron, desempenhando o papel de modelo substituto, aprenda durante o processo de otimização com algoritmo genético e passe a substituir, com uma probabilidade de substituição

cada vez maior, a função objetivo original. Na literatura, faz-se menção, também, a outros métodos, além do perceptron, capazes de desempenhar o papel de modelo substituto, tais como os métodos que empregam polinômios e os da geoestatística, segundo Naidu (2004). O uso do algoritmo genético se justifica pelo fato de se estar considerando que esta abordagem se aplica, especificamente, às funções objetivo cujo gradiente não é dedutível. Para demonstrar que este conceito é capaz de melhorar o desempenho da otimização com modelos substitutos, uma função multiobjetivo é utilizada como estudo de caso e os resultados são confrontados com a otimização sem o uso deste recurso. Para identificar a melhor abordagem, basta observar aquela que gera pontos mais próximos da fronteira de Pareto, em toda a sua extensão, em função da quantidade de avaliações da função objetivo original que se deseja realizar. Esta forma de identificar a melhor abordagem é suficiente, pois, no caso de uma função objetivo com custo computacional elevado, na qual esta abordagem se aplica, deseja-se que a convergência ocorra em determinado intervalo de tempo. Deste modo, a duração da otimização se torna um requisito e determina a quantidade de vezes que a função objetivo original pode ser avaliada, levando em consideração os meios computacionais disponíveis e o uso da computação paralela.

Qual é a razão para se supor que o uso de um modelo substituto torne o processo de otimização mais eficiente resultando na convergência mais rápida em função da quantidade de ativações da função objetivo original? A razão é a seguinte, partindo do pressuposto de que o modelo substituto é uma boa aproximação da função objetivo original, com baixo custo computacional, e que, para compensar o seu uso, o tempo gasto para prepará-lo é muito menor do que o tempo para realizar a otimização sem este recurso, ao invés do algoritmo de otimização avaliar a função objetivo original, ele avalia o modelo substituto, uma quantidade de vezes muito maior do que o modelo original, resultando na maior eficiência do processo de otimização.

O modelo substituto serve como um artifício para orientar o algoritmo de otimização, exigindo menos avaliações da função objetivo original. A estratégia mais simples de otimização com modelo substituto é a seguinte. Primeiro, avalia-se a função objetivo original, certa quantidade de vezes, a fim de gerar exemplos para construir o modelo substituto. Estes exemplos devem ser bem distribuídos no espaço de busca a fim de se obter uma boa representação da função objetivo original. Em seguida, atribui-se uma probabilidade de substituição e realiza-se a otimização. Por exemplo, se o probabilidade é de 0.9 e o algoritmo de otimização avalia a função objetivo 100 vezes, 90 avaliações são feitas pelo modelo substituto e 10 pela função objetivo original. Por último, parte do conjunto de soluções finais decorre da avaliação com o modelo substituto. Assim, para garantir a fidelidade dos resultados, toma-se esta parte e a submete à função objetivo original. Desta forma, a quantidade de vezes em que a função objetivo original é avaliada corresponde à quantidade de exemplos para construir o modelo substituto, à probabilidade de avaliações durante o processo de otimização, complementar ao de substituição, e à quantidade de soluções finais avaliadas pelo modelo substituto. Para justificar o uso do modelo substituto, a quantidade total de avaliações da função original, com este recurso, deve ser suficiente para se chegar ao resultado da otimização, em certo intervalo de tempo, imposto como requisito.

Há outras formas mais sofisticadas de otimização com modelo substituto. Por exemplo, imagine o processo de otimização realizado por ciclos. No final de cada ciclo, uma probabilidade das soluções pesquisadas decorre das avaliações do modelo original. Na

estratégia mais simples, descrita nos parágrafos anteriores, estas soluções são abandonadas. Ora, se elas ocorrem de qualquer forma, então por que razão não utilizá-las para melhorar o modelo substituto? Além disso, com a seguinte vantagem: à medida que o processo de otimização avança, estas soluções ocorrem cada vez mais próximas das regiões promissoras do espaço de busca, melhorando a precisão do modelo substituto nessas regiões.

O desafio desta técnica é construir modelos substitutos que representem com a maior precisão possível o modelo original, usando a menor quantidade possível de avaliações deste modelo, segundo Gorissen et al. (2010). Sob outro ponto de vista, trata-se de buscar um compromisso entre a alta fidelidade e a capacidade de realizar a otimização em determinado intervalo de tempo, de acordo com Osorio e Bierlaire (2009). Portanto, a proposta deste trabalho é atribuir uma probabilidade de substituição cada vez maior durante o processo de otimização com base no pressuposto de que o modelo substituto, a cada ciclo, pode ter sua precisão melhorada em função de novas soluções pesquisadas, cada vez mais próximas das regiões promissoras do espaço de busca, decorrentes das avaliações com a função objetivo original, durante o processo de otimização.

2. Estudo de Caso

O estudo da otimização com modelo substituto segue o roteiro abaixo.

- Passo 1: Defina a arquitetura e os parâmetros do perceptron.
- Passo 2: Defina os parâmetros do algoritmo genético.
- Passo 3: Para cada ensaio, execute os passos de 4 a 11.
- Passo 4: Gere o conjunto inicial de exemplos para o treinamento do perceptron.
- Passo 5: Para cada ciclo de otimização, execute os passos de 6 a 9.
- Passo 6: Execute o treinamento do perceptron.
- Passo 7: Execute o algoritmo genético com o recurso do modelo substituto.
- Passo 8: Adicione as soluções geradas pelas avaliações da função objetivo original ao conjunto de exemplos.
- Passo 9: Aumente a probabilidade de substituição da função objetivo original.
- Passo 10: Tome as soluções finais geradas pela avaliação do modelo substituto e volte a avaliá-las usando a função objetivo original.
- Passo 11: Execute o algoritmo genético sem o recurso do modelo substituto, realizando a mesma quantidade de avaliações da função objetivo original dos passos de 4 a 10.

Passo 12: Confronte as soluções finais de todos os ensaios, dos passos de 4 a 10 com os resultados do passo 11.

Passo 13: Classifique como a melhor abordagem aquela que gera pontos mais próximos da fronteira de Pareto.

Este estudo de caso consiste em minimizar a função abaixo (Equação 1), seguindo o roteiro acima.

$$F(X) = \begin{bmatrix} f_1 = (x_1 - 2)^2 + (x_2 - 1)^2 + 2 \\ f_2 = 9 \cdot x_1 - (x_2 - 1)^2 \end{bmatrix}, \quad -20 \leq x_i \leq 20, \quad i \in \{1, 2\} \quad (1)$$

Primeiro, definem-se a arquitetura e os parâmetros do perceptron (Tabela 1).

Tabela 1. Configuração do Perceptron

Parâmetro	Valor
Função de ativação	$\varphi(v) = \frac{1}{1 + e^{-2 \cdot v}}$
Quantidade de camadas escondidas	1
Quantidade de neurônios na camada escondida	10
Percentual de padrões de teste	0
Taxa de aprendizagem	0.1
Momento	0.5
Quantidade máxima de épocas de treinamento	3000
Varição mínima do erro quadrático médio	0.0001

Em seguida, definem-se os parâmetros do algoritmo genético (Tabela 2).

Tabela 2. Configuração do Algoritmo Genético

Parâmetro	Valor
Versão	MOGA Não-geracional
Quantidade de indivíduos (np)	70
Quantidade total de gerações (ng _{tot})	150

Por último, definem-se os demais parâmetros do estudo de caso (Tabela 3).

Tabela 3. Parâmetros do Estudo de Caso

Parâmetro	Valor
Quantidade de ensaios	10
Método de amostragem	LHS (uniforme)
Quantidade de exemplos iniciais (ns)	15
Quantidade de ciclos de otimização (N)	3
Probabilidade de substituição no 1º ciclo	0.94
Probabilidade de substituição no 2º ciclo	0.96
Probabilidade de substituição no 3º ciclo	1
Probabilidade de substituição média (P)	0.967

Este estudo compreende 10 ensaios (Tabela 3). Segue, abaixo, o resultado do primeiro ensaio.

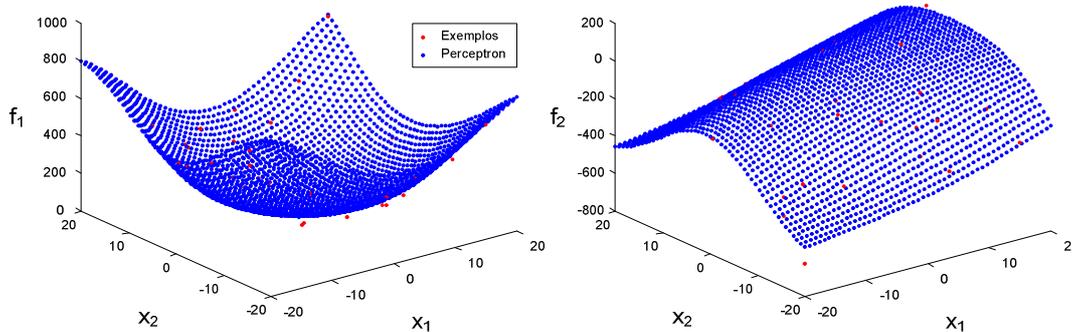


Figura 1. Modelo Substituto

No passo 6 do roteiro deste estudo, o perceptron assume a forma da função objetivo original (Figura 1). O uso do LHS (*Latin Hypercube Sampling*), como método de amostragem, garante uma boa distribuição dos exemplos no espaço de busca.

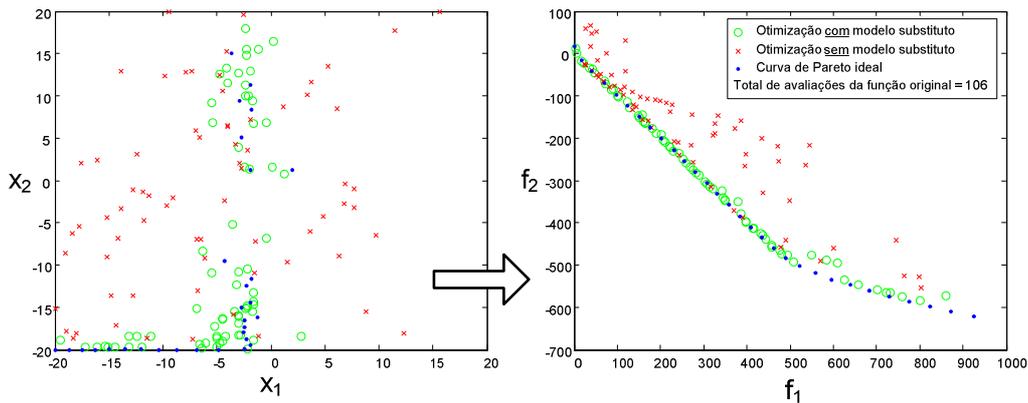


Figura 2. Resultado do Primeiro Ensaio

A Figura 2 mostra o resultado surpreendente da otimização com modelo substituto em comparação com o resultado da otimização sem este recurso, observando que na execução das duas abordagens o número de avaliações da função objetivo original é exatamente o mesmo. Porém, o que dizer de um estudo com 10 ensaios, considerando a sobreposição dos resultados de todos os ensaios? (Figura 3)

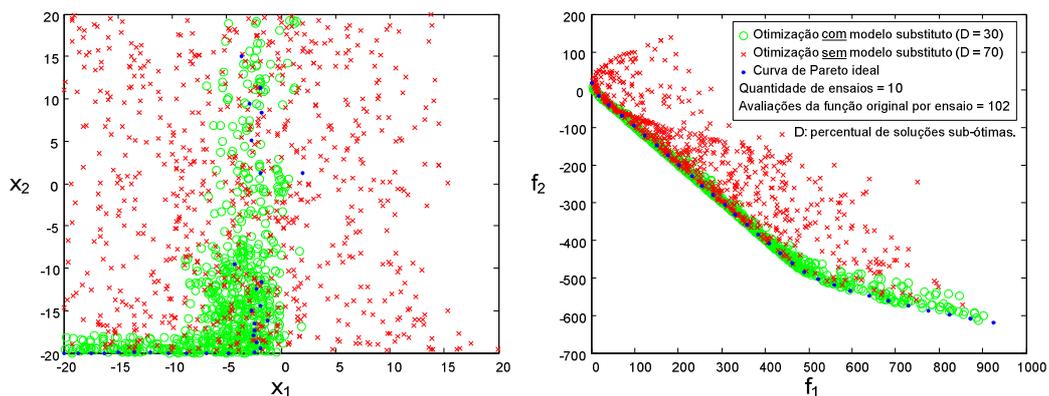


Figura 3. Resultado Final

A Figura 3 permite concluir que, neste estudo de caso, a otimização com modelo substituto teve um desempenho melhor do que a otimização sem o uso deste recurso, pois gerou pontos bem distribuídos e mais próximos da curva de Pareto ideal. Além disso, teve uma ocorrência menor de soluções sub-ótimas, apenas 30% do total de soluções, enquanto que a otimização sem este recurso teve 70% de soluções sub-ótimas.

Para explicar este resultado, elaborou-se o raciocínio abaixo.

Seja Q_1 a quantidade de ativações da função objetivo original na otimização com modelo substituto.

Seja Q_2 a quantidade máxima de gerações na otimização sem modelo substituto para que a quantidade de ativações da função objetivo original não seja maior do que Q_1 .

O algoritmo genético utilizado neste estudo de caso, versão de Borges (1999), avalia apenas dois filhos a cada geração a partir da recombinação de dois pais. Com base nesta informação, deduz-se a Equação 2 que fornece Q_2 .

$$Q_2 = \frac{(1-P) \cdot N}{2} \cdot np + (1-P) \cdot ng_{tot} + \frac{1}{2} \cdot ns \quad (2)$$

Q_2 : quantidade de gerações.

np : quantidade de indivíduos.

P : probabilidade média de substituição.

ng_{tot} : quantidade total de gerações.

N : quantidade de ciclos de otimização.

ns : quantidade de exemplos iniciais.

Seja Q_3 o percentual de soluções sub-ótimas na otimização com modelo substituto.

Seja Q_4 a quantidade mínima de gerações na otimização sem modelo substituto para que o percentual de soluções sub-ótimas não seja maior do que Q_3 .

$$\text{Seja: } \Delta ng = Q_4 - Q_2$$

Se Δng é um valor positivo e cresce com a quantidade de indivíduos (np), isto significa que a vantagem da otimização com modelo substituto sobre a otimização sem o uso deste recurso cresce, também, com np , servindo como um indicador do desempenho de uma abordagem em relação à outra.

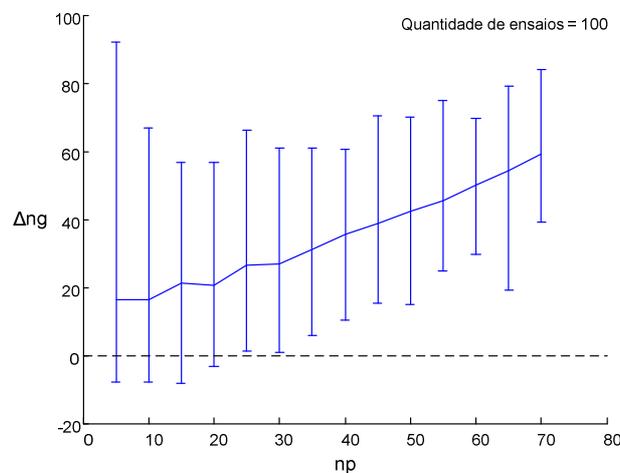


Figura 4. Desempenho

A Figura 4 mostra que a otimização com modelo substituto aumenta sua vantagem sobre a otimização sem o uso deste recurso na medida em que se define uma quantidade cada vez maior de indivíduos na população, o que é conveniente no caso de uma otimização multiobjetivo, pois uma quantidade maior de indivíduos ajuda a caracterizar melhor a fronteira de Pareto em toda sua extensão.

Na Figura 4, a linha contínua conecta as médias de Δng e as barras de erro são os máximos e mínimos alcançados durante a realização de 100 ensaios.

3. Conclusão

O objetivo deste trabalho foi investigar se a otimização com modelo substituto, executada por ciclos e com probabilidade de substituição crescente, durante o processo de otimização, tinha um desempenho melhor do que a otimização sem uso deste recurso, quando aplicada a um problema onde a função objetivo tem um custo computacional elevado e o seu gradiente não é dedutível. Para alcançar este objetivo, realizou-se um estudo de caso com uma função multiobjetivo, realizando 10 ensaios. Os resultados mostraram um desempenho surpreendente da otimização com modelo substituto em relação à otimização sem o uso deste recurso. Para explicar este resultado, deduziu-se uma forma de quantificar o desempenho de uma abordagem em relação à outra. Concluiu-se o seguinte: a otimização com modelo substituto leva vantagem em relação à otimização sem este recurso quando se define uma quantidade maior de indivíduos; e, a quantidade maior de indivíduos é vantajosa quando o problema é multiobjetivo, pois isto ajuda a caracterizar melhor a fronteira de Pareto.

Referências

- BORGES, C. C. H. (1999) Algoritmos Genéticos para Otimização em Dinâmica de Estruturas. 216 p. Universidade Federal do Rio de Janeiro (UFRJ).
- GORISSEN, D.; COUCKUYT, I.; DEMEESTER, P.; DHAENE, T.; CROMBECQ, K. A Surrogate Modeling and Adaptive Sampling Toolbox for Computer Based Design. *Journal of Machine Learning Research*, v. 11, p. 2051-2055, July 2010. Disponível em: <<http://jmlr.csail.mit.edu/papers/volume11/gorissen10a/gorissen10a.pdf>>. Acesso em: 20 set. 2010.
- NAIDU, S. L. S. (2004) Neural Network Surrogate Model For Multidisciplinary Design Optimization. 100 p. Dissertation (Master of Technology) – India Institute of Technology, Bombay.
- OSORIO, C.; BIERLAIRE, M. A surrogate model for traffic optimization of congested networks: an analytic queueing network approach. Ecole Polytechnique Fédérale de Lausanne, 2009. Disponível em: <<http://transp-or2.epfl.ch/technicalReports/osorBier09.pdf>>. Acesso em: 20 set. 2010.