

Assimilação de dados com redes neurais artificiais

Helaine C. M. Furtado¹, Haroldo F. de Campos Velho², Elbert A. N. Macau²

¹Programa Doutorado em Computação Aplicada – CAP
Instituto Nacional de Pesquisas Espaciais – INPE

²Laboratório Associado de Computação e Matemática Aplicada – LAC
Instituto Nacional de Pesquisas Espaciais – INPE

{helaine.furtado,haroldo,elbert, etc}@email.br

Abstract. *Description of a physical phenomenon through differential equations has errors involved, since the mathematical model is always an approximation of reality. For an operational prediction system, one strategy to improve the prediction is to add some information from the real dynamics into mathematical model. This additional information consists of observations on the phenomenon. However, the observational data insertion should be done carefully, for avoiding a worse performance of the prediction. Technical data assimilation are tools to combine data from physical-mathematics model with observational data to obtain a better forecast.*

Resumo. *Realizar previsão de fenômenos físicos por meio de modelagem matemática é uma atividade passível de erro, uma vez que, o modelo é sempre uma aproximação da realidade. Uma alternativa para “melhorar” a descrição do modelo é fazer uso de medidas do sistema em estudo. No entanto, as medidas devem ser cuidadosamente inseridas no modelo para evitar piora na execução da previsão. Técnicas de assimilação de dados são uma ferramenta que combinam de modo eficiente dados de observação com dados de um modelo físico-matemático para obter uma melhora na previsão do modelo.*

Palavras-chave: *assimilação de dados, redes neurais artificiais, equação da onda, sistema de Lorenz*

1. Introdução

Descrever fenômenos físicos a partir de equações diferenciais (parcial ou ordinária), é uma atividade passível de erro, uma vez que o modelo matemático é sempre uma aproximação da realidade. Um modo de “melhorar” esta descrição é adicionando informações ao modelo. Estas informações consistem de observações, que são dados provenientes de medidas sobre o fenômeno em estudo. No entanto, ao usar observações para estimar alguma propriedade do modelo matemático (condições iniciais, condições de contorno, propriedades do material e/ou da dinâmica, termos fontes) o problema de estimação é frequentemente um problema mal posto [Hadamard 1952]. A alternativa para a solução do problema é combinar de modo eficiente as observações com o modelo matemático, por meio da utilização de técnicas de assimilação de dados, que podem ser baseadas em teoria da estimação (como: filtro de Kalman [Kalman 1960], filtro de partículas [Gordon et al. 1993, Chorin and Krause 2004]); baseadas no cálculo variacional [Talagrand and Courtier 1987, Bennett 2002]; ou ainda, baseadas em inteligência artificial, com o uso de Redes Neurais Artificiais [Nowosad et al. 2000].

Com o desenvolvimento do sistema de observação e a evolução dos modelos numéricos, a quantidade de dados disponíveis para serem assimilados é da ordem de 10^5 a 10^7 , enquanto que, o número de pontos de grade do modelo é da ordem de 10^6 a 10^8 . Isso representa uma demanda de computação intensiva. É um grande desafio desenvolver técnicas que possam produzir um resultado de análise o mais preciso possível para produzir boa capacidade de previsão, que seja mais eficiente e tornando o custo de computação operacionalmente viável, em um contexto em que, o número de pontos de grades e de observações são exponencialmente crescentes. Assim, este trabalho tem como objetivo apresentar a ferramenta de redes neurais artificiais como uma técnica de assimilação de dados.

2. Aplicações

Os modelos de aplicações utilizados para testar a metodologia foram o sistema de Lorenz [Lorenz 1963], que é um sistemas de equações diferenciais parciais e o modelo de onda dado pela equação 1.

$$\frac{\partial \eta_F}{\partial t} + c \frac{\partial \eta_F}{\partial x} = F(x, t) \quad (1)$$

em que η é o deslocamento, c é a constante de velocidade de fase, F é uma forçante externa, t é o tempo e x é o espaço. O *subscript* F indica a solução avançada, isto é, a solução *a priori*. A Equação 1 pode ser interpretada também como a equação da convecção, em que η é a concentração e c é o coeficiente de convecção.

3. Filtro de Kalman

A partir de um modelo de previsão e um sistema de observação, em que o subscrito n denota o tempo discreto, o sobrescrito p representa o valor previsto, tem-se:

$$\begin{aligned} \mathbf{x}_{n+1}^p &= \mathbf{F}_n \mathbf{x}_n^p + \mu_n \\ \mathbf{y}_n^p &= \mathbf{H}_n \mathbf{x}_n^p + \nu_n \end{aligned} \quad (2)$$

em que $\mathbf{F}_n$ representa o modelo matemático, μ_n é uma forçante estocástica (ruído aleatório do modelo). O sistema de observações é modelado por um operador $\mathbf{H}_n$ e ν_n é o ruído associado as observações. A hipótese para a função densidade de probabilidade Gaussiana com média zero é utilizada para os ruídos de modelagem e de observação, $\mathbf{y}_n^p$ é o vetor de observações. O vetor de estado $\mathbf{x}_{n+1}$ é estimado recursivamente no tempo por meio de:

$$\mathbf{x}_{n+1}^a = (\mathbf{I} - \mathbf{G}_{n+1} \mathbf{H}_{n+1}) \mathbf{F}_n \mathbf{x}_n^a + \mathbf{G}_{n+1} \mathbf{y}_{n+1}^p \quad (3)$$

sendo que $\mathbf{x}_{n+1}^a$ é o valor de análise estimado, $\mathbf{G}_n$ é a matriz de ganho de Kalman, que é calculada da a partir da minimização variância do erro J_{n+1} [Jazwinski 1970]: O método de Filtro de Kalman como uma técnica de assimilação de dados pode ser encontrada [Furtado 2008].

4. Redes Neurais Artificiais

Redes neurais artificiais (RNA) tem se tornado uma importante ferramenta em processamento de sinais. Muitas pesquisas têm sido realizadas na busca de novos modelos de RNA adaptando o modelo existente para resolver problemas da vida real, tais como os problemas de engenharia [Haykin 2004]. O funcionamento das RNA é realizado por meio

de duas fases: a fase de treinamento¹ e a fase de ativação². O algoritmo usado para o treinamento do PMC - Perceptron de Múltiplas Camadas foi o *backpropagation*, a qual é uma rede *feedforward* composta por uma camada de entrada, uma camada de saída e camadas escondidas, cujo o objetivo é extrair propriedades estatísticas da entrada dos dados [Haykin 2004]. O método de assimilação de dados baseado em RNA, por exemplo, o Perceptron Multicamadas pode ser escrito como:

$$\mathbf{x}_n^a = F_{RNA}(\mathbf{x}_n^p, \mathbf{y}_n^o, \mathbf{x}_n^{FK}) \quad (4)$$

em que $\mathbf{x}_n^a$ é a análise, F_{RNA} representa o processo de assimilação por RNA, $\mathbf{x}_n^p$ são os dados do modelo, $\mathbf{y}_n^o$ os dados de observação e $\mathbf{x}_n^{FK}$ a análise obtida com o Filtro de Kalman descrito na seção 3, que é usado na fase de treinamento da rede.

5. Resultados e discussões

Para a implementação do Filtro de Kalman, foram usadas as seguintes matrizes: $\mathbf{Q}_t = 0.1\mathbf{I}$; $\mathbf{R}_t = 0,5\mathbf{I}$; $\mathbf{H} = \mathbf{I}$; matriz de covariância do erro de modelagem, matriz de covariância do erro de observação e o operador que representa o sistema de observação, respectivamente. A matriz de covariância do erro de previsão é inicializada por:

$$\mathbf{P}_0^f = \begin{cases} 10(x_0^f)^2 & \text{para } i = j, \\ 0 & \text{para } i \neq j. \end{cases} \quad (5)$$

As observações assimiladas neste trabalho foram geradas artificialmente. Os dados foram gerados a partir da integração do sistema, adicionando-se um ruído aleatório de variância 0,5. A referência de verdade para o método de assimilação é a curva obtida a partir da integração dos modelos sem ruído, o que significa que, quanto mais próximo as estimativas estiverem da dinâmica do modelo, melhor será a estimativa obtida pelos métodos de assimilação.

A rede Perceptron de Múltiplas Camadas, descrita na seção 4, é uma rede que exige o treinamento supervisionado, isto é, que seja apresentado a ela a saída desejada, ou o denominado “alvo” da rede. A saída desejada da rede é a estimativa obtida com o Filtro de Kalman, descrito da seção 3. A fase de treinamento da rede, consiste em determinar os melhores pesos que mapeia os dados de entrada aos dados da saída desejada (“alvo” da rede). A rede implementada possui uma camada escondida com dois neurônios; dois neurônios na camada de entrada e um neurônio na camada de saída. Foram utilizados 4800 padrões de treinamento, o critério de parada foi o número de épocas de treinamento, neste experimento foram 100 épocas de treinamento.

O modelo de onda linear foi integrado com o método de Crank-Nicholson. As condições iniciais e de contorno são periódicas. A condição inicial usada, foi a solução analítica da equação de Korteweg-de Vries. Na Tabela 1 apresenta-se os resultados dos erros de assimilação, que é a média da diferença absoluta entre a verdade e o estimado.

O método das redes neurais artificiais para ambos os modelos foi mais eficiente que o Filtro de Kalman, pode-se observar pelo o valor do erro apresentado na Tabela 1. Além disso, o tempo de execução do FK foi de 5 minutos e com a rede neural foi 0,83 minutos, menos de um minuto. Ambas as técnicas foram implementadas no Matlab e executadas no computador *Intel Core 2 Duo*, com 2GB de memória ram.

¹Um conjunto de exemplos são apresentados à rede, a qual extrai as características necessárias para representar a informação fornecida.

²Utiliza as características armazenadas nos pesos após a fase de treinamento.

Tabela 1. Resultados do erro de assimilação

	FK	RNA-PMC
Lorenz	1,2510	0,3941
onda1D	1,3178	0,1278

Referências

- Bennett, A. F. (2002). *Inverse Modeling of the Ocean and Atmosphere*. Cambridge University Press, New York, 1 edition.
- Chorin, A. J. and Krause, P. (2004). Dimensional reduction for a bayesian filter. *PNAS*, 101(42):15013–15017.
- Furtado, H. C. M. (2008). Redes neurais e diferentes métodos de assimilação de dados em dinâmica não linear. Master’s thesis, Instituto Nacional de Pesquisas Espaciais, São José dos Campos.
- Gordon, N. J., Salmond, D., and Smith, A. (1993). Novel approach to nonlinear/non-gaussian bayesian state estimation. *IEE Proceedings*, 140(2):107–113.
- Hadamard, J. (1952). J.(1952) lectures on cauchy’s problem.
- Haykin, S. (2004). *Redes neurais princípios e prática*, volume 2. Editora Bookman, Porto Alegre.
- Jazwinski, A. H. (1970). *Stochastic processes and filtering theory*. New York, NY: Academic, New York.
- Kalman, R. E. (1960). A new approach to linear filtering and prediction problems. *Journal of Basic Engineering*, (82):35–45.
- Lorenz, E. N. (1963). Deterministic nonperiodic flow. *Journal of the Atmospheric Sciences*, 20(2):130–141.
- Nowosad, A., Campos Velho, H., and Rios Neto, A. (2000). Neural network as a new approach for data assimilation. In *Congresso Brasileiro de Meteorologia*, volume 1, pages 3078–3086, Rio de Janeiro, (Brasil). Proceedings in CD-ROM (paper code PT00002).
- Talagrand, O. and Courtier, P. (1987). Variational assimilation of meteorological observations with the adjoint vorticity equation. *Quart. J. Roy. Meteor. Soc.*, 113:1311–1328.