MÉTODOS HÍBRIDOS KALMAN/VAR PARA ASSIMILAÇÃO DE DADOS: CONCEITOS PRELIMINARES

BALBONI, Mauricio Dorneles Caldeira[1]; BECK, Vinicius Carvalho[2].

[1] UFPEL - baalbis@gmail.com

[2] UAB/UFPEL - vonoco@gmail.com

RESUMO

A assimilação de dados é uma metodologia de combinar modelos matemáticos e dados de observação, que vem a ser uma alternativa para a simples interpolação quando o número de graus de liberdade do modelo numérico utilizado é bem maior do que o número de dados disponíveis. Existem várias técnicas para assimilação, sendo que as principais são a filtragem de Kalman e a assimilação variacional. Recentemente na literatura, surgiram pesquisas envolvendo a utilização de métodos híbridos, combinando filtragem de Kalman com métodos variacionais. O objetivo deste trabalho é apresentar alguns detalhes matemáticos destes métodos híbridos, particularmente com relação às combinações lineares utilizadas nestes métodos.

Palavras-chave: assimilação de dados, Kalman, VAR.

1. INTRODUÇÃO

O procedimento de combinar modelos matemáticos e dados de observação, que vem a ser uma alternativa para a simples interpolação quando o número de graus de liberdade de modelos numéricos é superior ao número de dados disponíveis. Atualmente, técnicas de assimilação são amplamente utilizadas em áreas tais como meteorologia (redução de erro dos modelos numéricos), robótica (correção de rotas) e engenharia de satélites (correção de órbitas).

Existem várias técnicas para assimilação, sendo que as principais são a filtragem de Kalman e a assimilação variacional. Desde as primeiras utilizações destes métodos, muitas variações de cada metodologia já foram introduzidas. Embora existam vários estudos comparando os dois métodos básicos e seus derivados (MENG; ZHANG, 2008; ZHANG et al., 2011), os resultados dos testes realizados mostram que, em geral, não se pode afirmar qual metodologia é mais eficiente, tanto do ponto de vista matemático quanto computacional. Por esta razão, surgiram recentemente na literatura pesquisas envolvendo a utilização de métodos híbridos, combinando filtragem de Kalman com métodos variacionais.

HAMILL; SNYDER (2000) desenvolveram uma forma híbrida entre o 3DVAR e o FKEns (uma variação do FK), onde a matriz de covariância dos erros de estimativa é obtida através da combinação linear convexa das matrizes de covariância dos erros calculadas com o método 3DVAR e com o FKEns. Os autores concluíram que as estimativas mais precisas para a maioria das variáveis foram obtidas com maior peso atribuído ao FKEns. HANSEN; SMITH (2001) propuseram uma análise semelhante baseada na integração do 4DVAR (extensão do 3DVAR) e do FKEns paralelamente, onde a matriz de covariância do erro de estimativa do FKEns foi utilizada na minimização da função custo do 4DVAR.

WANG et al. (2008) propuseram um método híbrido, utilizando o FKTEns e o 3DVAR, onde as matrizes de covariância dos erros, geradas pelos *ensembles* do FKEns, são utilizadas na minimização variacional.

O objetivo deste trabalho é apresentar alguns detalhes matemáticos destes métodos híbridos, particularmente com relação às combinações lineares utilizadas no cálculo da matriz da covariância dos erros de estimativa dos modelos.

2. MATERIAL E MÉTODOS

Nas seções 2.1 e 2.2 são apresentados o Método Variacional Tridimensional (3DVAR) e o Filtro de Kalman (FK), respectivamente. Na seção 2.3 é apresentada a definição de combinação linear convexa. A relação desta definição com a forma híbrida dos métodos citados será introduzida e discutida na seção de resultados e discussão.

2.1 Método 3DVAR

O método 3DVAR consiste, basicamente, na minimização de uma função custo J, diretamente proporcional à diferença entre observação e estimativa do modelo, com o objetivo de se obter a análise ótima. A idéia de representar o erro através de uma função custo teve origem no trabalho de SASAKI (1958).

A função custo descrita por LORENC (1986), com base na teoria de probabilidades bayesiana, é dada por

$$J(x) = \frac{1}{2}(x - x_b)^T B^{-1}(x - x_b) + \frac{1}{2}(y_0 - H(x))^T R^{-1}(y_0 - H(x)) \quad \text{Equação 1}$$

onde x é a observação, x_b é o campo de *background* (integração curta do modelo ou climatologia, que pode ser conotado como campo suporte), y_0 é a estimativa inicial, H é um operador que age sobre a dimensão da observação possibilitando a comparação desta com a estimativa inicial do modelo, B é a matriz de covariância dos erros de estimativa e R é a matriz de covariância dos erros de observação.

A expressão apresentada na equação 2 é a solução exata para o problema variacional 3DVAR de minimização da função custo $J(x)$.

$$x_a = x_b + (B^{-1} + H^T R^{-1} H)^{-1} H^T R^{-1}(y_0 - H(x_b)). \quad \text{Equação 2}$$

2.2 Filtro de Kalman

A seguir, será descrito, resumidamente a técnica FK (KALMAN, 1960). Um detalhamento matemático mais completo do FK pode ser encontrado em GONÇALVES (2005). Após uma estimativa inicial para w_0^f (campo previsto) e P_0^f (matriz de covariância do erro de estimativa), utilizando o sobreíndice f - do inglês *forecasting* - para significar fase de propagação ou previsão, e o sobreíndice a para significar fase de atualização ou assimilação, cada iteração de uma aplicação do FK passa por quatro etapas:

1) Previsão a partir do modelo:

$$w_{i+1}^f = F_i w_i^a \quad \text{Equação 3}$$

$$P_{i+1}^f = F_i P_i^a F_i^T + Q_i \quad \text{Equação 4}$$

onde w_{i+1}^f é a previsão, F_i é a matriz de dinâmica do sistema, w_i^a é a análise, P^f é a matriz de covariância dos erros de estimativa, P^a é a matriz de covariância dos erros de estimativa obtida pela análise e Q_i é a matriz de covariância do ruído de observação.

2) Cálculo da Matriz ganho:

$$G_{i+1} = P_{i+1}^{f} H_{i+1}^{T} [R_{i+1} + H_{i+1} P_{i+1}^{f} H_{i+1}^{T}]^{-1} \quad \text{Equação 5}$$

onde G_{i+1} é a matriz ganho, H_{i+1} é um operador que transforma as grandezas medidas pelos instrumentos meteorológicos nas grandezas utilizadas pelo modelo e R_{i+1} é a matriz de covariância dos erros de observação.

3) Cálculo da estimativa:

$$w_{i+1}^{est} = H_{i+1} w_{i+1}^{f} \quad \text{Equação 6}$$

onde w_{i+1}^{est} é uma estimativa para o vetor w_{i+1} real (sempre desconhecido).

4) Análise:

$$w_{i+1}^{a} = w_{i+1}^{f} + G_{i+1}(w_{i+1}^{obs} - w_{i+1}^{est}) \quad \text{Equação 7}$$

onde w_{i+1}^{obs} é o vetor w_{i+1} medido (com erro de medição).

$$P_{i+1}^{a} = [I - G_{i+1} H_{i+1}] P_{i+1}^{f} \quad \text{Equação 8}$$

onde I é a matriz identidade.

2.3 Combinações Lineares Convexas

Seja E um espaço vetorial, com $u, v \in E$. Chama-se de segmento de reta de extremidades u e v, o conjunto $[u, v] = \{(1-\alpha)u + \alpha v; 0 \leq \alpha \leq 1\}$. Um conjunto X chama-se convexo, quando $u, v \in X \Rightarrow [u, v] \subset X$, ou seja, quando o segmento de reta que une dois quaisquer de seus pontos ainda está contido em X (LIMA, 1998). Uma combinação linear convexa de dois vetores (em particular, de duas matrizes) $u, v \in E$, é um vetor (ou matriz) da forma $(1-\alpha)u + \alpha v$, com $\alpha \in \mathbb{R}$.

3. RESULTADOS E DISCUSSÃO

No híbrido FK/3DVAR, por exemplo, utiliza-se a matriz de covariância $P_{i+1}^{f} = (1-\alpha)P_i^{f} + \alpha B_i$, com $\alpha \in \mathbb{R}$, com P_i^{f} do FK, e B_i do método 3DVAR. Esta matriz híbrida pode ser utilizada tanto no FK, quanto no 3DVAR, ou em ambos paralelamente.

Ressalta-se que este é apenas o primeiro passo teórico para o entendimento dos métodos híbridos em assimilação de dados. A partir deste estudo inicial, pode-se introduzir uma sistematização dos métodos híbridos, a qual poderá servir como embasamento para testes preliminares com métodos híbridos.

Um amplo campo de estudo poderá surgir, tendo em vista a grande quantidade de variações dos métodos baseados em filtragem de Kalman e variacionais. Em pesquisas futuras, pretende-se testar diferentes configurações de métodos híbridos em modelos de equações primitivas, como o modelo de Lorenz, por exemplo.

4. CONCLUSÃO

Os métodos híbridos de assimilação de dados constituem uma alternativa a ser testada, em um primeiro momento, nos modelos de equações primitivas. Existe uma grande quantidade de versões de cada método, possibilitando várias combinações possíveis. Este trabalho apresenta um estudo teórico inicial destes métodos, possibilitando uma continuação, tanto no âmbito teórico, ampliando o entendimento matemático de cada método, quanto no âmbito aplicado, testando implementações dos métodos discutidos teoricamente.

REFERÊNCIAS BILIOGRÁFICAS

GONÇALVES, Dimas José. **Aspectos matemáticos do Filtro de Kalman Discreto.** 2005. 54f. Dissertação (Mestrado em Matemática) – Departamento de Matemática, Estatística e Computação Científica, Universidade Estadual de Campinas.

HAMILL, Thomas M.; SNYDER, Chris. A hybrid ensemble Kalman filter - 3D variational analysis scheme. **Monthly Weather Review**, v. 128, n. 8, p. 2905-2919, 2000.

HANSEN, James A.; SMITH, Leonard A. Probabilistic noise reduction. **Tellus**, V. 53, n. 5, p. 585-598, 2001.

KALMAN, R. E. A new approach to linear filtering and prediction problems, **Journal Basic Engineering** , v. 82D, p. 35-45, 1960.

LIMA, Elon Lages. **Álgebra Linear**. 3º edição. Instituto Nacional de Matemática Pura e Aplicada, Rio de Janeiro, 1998.

LORENC, A. C. Analysis methods for numerical weather prediction. **Quarterly Journal of Royal Meteorology Society**, v. 112, n. 474, p. 1177-1194, 1986.

MENG, Z.; ZHANG, F. Tests of an ensemble Kalman filter for Mesoscale and regional-scale data assimilation. Part III: Comparison with 3DVAR in a real-data case study. **Monthly Weather Review**, v. 136, p. 522-540, 2008.

SASAKI, Y. An objective analysis based on the variational method. **Journal of Meteorological Society of Japan**, v. 36, n. 3, p. 77-88, 1958.

WANG, X.; BARKER, D. M.; SNYDER, C.; HAMILL, T. M. A hybrid ETKF-3DVAR data assimilation scheme for the WRF model. Part I: observing system simulation experiment. **Monthly Weather Review**, v. 136, p. 5116-5131, 2008.

ZHANG, M.; ZHANG, F.; HUANG, X. -Y.; ZHANG, X. Intercomparison of an Ensemble Kalman Filter with Three- and Four- Dimension Variational Data Assimilation Methods in a Limited-Area Model over the Month of June 2003. **Monthly Weather Review**, v. 139, p. 566-572, 2011.