

## GRADIENTES E O DESEMPENHO DE MODELOS NUMÉRICOS

Paulo Sérgio Dias da Silva<sup>1</sup> e Luiz Landau<sup>2</sup>

<sup>1</sup> Universidade Estadual do Norte Fluminense  
Laboratório de Ciências Matemáticas  
Av. Alberto Lamego, 2000  
28013-600-Campos dos Goytacazes-RJ-Brasil  
e-mail :[paulosd@uenf.br](mailto:paulosd@uenf.br)

<sup>2</sup> Universidade Federal do Rio de Janeiro  
COPPE -Programa de Engenharia Civil  
Caixa Postal 68552  
21949-900-Rio de Janeiro-RJ-Brasil  
e-mail:[landau@lamce.ufrj.br](mailto:landau@lamce.ufrj.br)

### ABSTRACT

Given a differentiable function  $f : V \subset \mathbb{R}^n \rightarrow \mathbb{R}$ , the properties of the gradient vector of  $f$  are well known for each  $x \in V$ . It will be examined a non trivial problem in the domain of the Applied Mathematics in which the role of the gradient of a certain functional in the definition of the differential of this functional is essential to the development of an efficient algorithm in order to obtain one configuration of the evolutionary system, when we have a numerical model of the system dynamics and a space-time distributed set of observations of the same system, useful in the prediction of the future behavior of the system under study, and which can also be used as a powerful method for the analysis of the numerical model performance.

**Keywords:** 4D variational data assimilation, control theory, applied mathematics, pollutant spillage detection.

### RESUMO

Dada uma função diferenciável  $f : V \subset \mathbb{R}^n \rightarrow \mathbb{R}$ , são bem conhecidas as propriedades do vetor gradiente de  $f$  definido para cada  $x \in V$ . Examinar-se-á um problema nada trivial da Matemática Aplicada em que o papel desempenhado pelo gradiente de um funcional na definição deste funcional é determinante para o desenvolvimento de um algoritmo eficiente à obtenção de uma configuração de um sistema evolutivo, conhecidos um modelo numérico da dinâmica do sistema e um conjunto de observações do mesmo sistema distribuídas tanto no espaço quanto no tempo, útil na predição do comportamento futuro do sistema em estudo, e que pode também ser empregado como um poderoso método de análise do desempenho do modelo numérico desenvolvido.

**Palavras-chave:** assimilação variacional de dados 4D, teoria do controle, matemática aplicada, determinação de derrame de poluentes.

### 1 - INTRODUÇÃO

Sejam  $V$  um aberto em  $\mathbb{R}^n$  e  $f : V \rightarrow \mathbb{R}$  uma função diferenciável. Então,  $\forall x \in V$ , a diferencial de  $f$  em  $x$  é o

funcional linear tal que,  $\forall v \in \mathbb{R}^n$ ,  
$$df(x) \cdot v = \lim_{t \rightarrow 0} \frac{f(x+tv) - f(x)}{t}.$$

Como  $df(x)$  é um funcional linear de  $(\mathbb{R}^n)^*$ , o espaço vetorial dual de  $\mathbb{R}^n$ , decorrem de um resultado básico da

Álgebra Linear a existência e a unicidade do vetor do  $\mathbb{R}^n$ ,  $\nabla f(x)$ , tal que

$$df(x) \cdot v = \langle \nabla f(x), v \rangle, \quad (1)$$

sendo  $\langle, \rangle$  o produto interno canônico de  $\mathbb{R}^n$ . Da expressão (1), decorrem as propriedades mais importantes do vetor  $\nabla f(x)$ , a saber [1]: 1.  $\nabla f(x)$  indica uma direção de crescimento da função  $f$  em  $x$ , 2. de todas as direções em  $\mathbb{R}^n$ ,  $\nabla f(x)$  é aquela em que o crescimento da

função é mais rápido e  $\nabla f(x)$  é normal à superfície de nível de  $f$  passando por  $x$ .

Recentemente, com a necessidade de se desenvolverem métodos numéricos eficientes ao tratamento de problemas de larga escala, a unicidade da representação (1) revelou outra propriedade importantíssima, ao permitir a resolução numérica de problemas exigindo um esforço muito acima da capacidade dos mais sofisticados computadores disponíveis até o presente momento.

Trata-se nos parágrafos seguintes de uma abordagem para problemas envolvendo sistemas dinâmicos de grande complexidade, dos problemas matemáticos e computacionais que essa abordagem levantou, do método que tornou tal abordagem computacionalmente tratável, permitindo sua aplicação a muitos outros problemas e, por sua vez, possibilitando novas áreas de investigação bastante promissoras.

## 2 - ASSIMILAÇÃO VARIACIONAL DE DADOS 4D

Considere-se a seguinte situação: estuda-se um sistema evolutivo  $S$  e, para tanto, dispõe-se de um conjunto de observações do sistema coletadas no intervalo de tempo  $[t_0, t_1]$  e distribuídas no espaço e no tempo, ou seja, a rede de instrumentos para a obtenção de dados sobre o sistema está espalhada pela região espacial onde reside o foco de interesse sobre  $S$  e coligindo informações durante o intervalo de tempo arbitrado, embora nem todos os instrumentos façam, ao mesmo tempo, seus registros. Além disso, tem-se um modelo numérico capaz de simular a dinâmica do sistema.

O objetivo é, com base nessas informações, determinar o estado ou a configuração de  $S$  num instante de tempo  $t_f > t_1$ , processo conhecido como Assimilação de Dados, entre aqueles que estudam os problemas relacionados com a dinâmica de fluxos geofísicos.

Os instrumentos empregados na coleta de informações sobre  $S$ , por mais sofisticados que sejam, contêm erros em seus registros. Contudo, são conhecidas as principais estatísticas sobre tais erros. Os recursos computacionais disponíveis, ainda que excelentes, operam com uma inevitável profusão de hipóteses assumidas no projeto do modelo numérico, gerando discrepâncias entre os estados do sistema produzidos pelo modelo e os reais estados de  $S$ , o que deve ser possível identificar e, ao menos, atenuar.

Uma idéia para se resolver o problema é combinar todas as informações que se tem sobre  $S$ , tanto as provenientes de suas observações quanto as fornecidas pelo conhecimento de sua dinâmica, para conseguir-se uma configuração de  $S$  que as concilie, de modo que os estados do sistema produzidos pelo modelo estejam em acordo com as observações conhecidas.

Operacionalmente, há duas classes de algoritmos capazes de realizar a Assimilação de Dados: seqüenciais e

variacionais. Nos algoritmos seqüenciais, a cada instante de tempo em que há observações disponíveis, estas são utilizadas para corrigir a predição do modelo para aquele instante (o que significa o estado do sistema gerado pelo modelo em um instante de tempo anterior ao considerado), produzindo um estado do sistema a partir do qual será feita uma nova predição do estado do sistema. Este método está intimamente relacionado com o formalismo conhecido como filtro de Kalman.

Nos algoritmos variacionais, busca-se um estado do sistema num instante de tempo, por exemplo,  $t_0$ , tal que, se  $X_0 = X(t_0)$  é a configuração do sistema no instante considerado, então os estados do sistema nos instantes de tempo posteriores a  $t_0$  e produzidos pelo modelo a partir de  $X_0$  estejam próximos, num sentido a ser bem definido posteriormente, das observações do sistema nos instantes de tempo correspondentes. Este último método de resolução do problema é conhecido como Assimilação Variacional de Dados 4D. É claramente iterativo, e nele, ao contrário dos algoritmos seqüenciais, as informações sobre o sistema em estudo propagam-se para a frente e para trás no tempo, visando um ajuste global da evolução do modelo aos dados observados.

Neste trabalho, adotou-se propositalmente um jargão não usual no meio estritamente matemático, embora a ênfase esteja na correta definição dos conceitos matemáticos, necessária à evolução das idéias e dos métodos científicos, isto porque o método da Assimilação Variacional de Dados 4D tem-se revelado adequado ao tratamento de uma vasta gama de situações da matemática aplicada e das engenharias. O problema de ajustar a trajetória de um corpo em movimento a um conjunto de observações deste corpo não é recente, ao contrário, basta recordar o árduo trabalho de Johannes Kepler para encontrar uma curva que se adaptasse às observações de Marte anteriormente coligidas por Tycho Brahe, ou os trabalhos de Gauss desde a determinação da órbita de Ceres até o desenvolvimento do Princípio dos Mínimos Quadrados [2]. Como sempre ocorre em matemática, tais trabalhos pioneiros originaram, ao longo de séculos, uma formidável estrutura teórica que, neste caso, envolveu, restringindo-se aos últimos cem anos aproximadamente, Henri Poincaré [3], Lev Pontriaguin [4], Rudolf Kalman [5], Jacques Louis Lions [6], dentre outros. Ainda assim, até meados da década de 1970, não era possível a aplicação direta dos métodos da Teoria do Controle a problemas envolvendo a dinâmica de fluxos geofísicos, uma vez que o número de variáveis nestes problemas tornava-os computacionalmente intratáveis, mesmo para os equipamentos atuais.

Foi somente o emprego do formalismo adjunto [7] que permitiu a drástica redução no número de variáveis de controle, inicialmente em problemas de circulação atmosférica, originando o método da Assimilação Variacional de Dados 4D.

Mas, pode-se perguntar, por que era tão importante determinar-se com tamanha precisão a configuração inicial do sistema em um problema de circulação atmosférica a ponto de ter-se de se lançar mão de um procedimento computacionalmente tão dispendioso como a Assimilação de Dados? A resposta a essa questão foi a descoberta quase acidental por Edward Norton Lorenz, por volta de 1962, do comportamento caótico da atmosfera terrestre. Tal descoberta, como bem apontou Olivier Talagrand [8], levou a comunidade de Previsão Numérica do Tempo, num primeiro momento, a crer que todo seu trabalho estaria destinado ao erro, exceto se limitado a um curtíssimo intervalo de tempo, dada a impossibilidade àquela época de se obter uma configuração da atmosfera terrestre isenta de erros. Logo, porém, teriam início os esforços para o desenvolvimento de métodos que sanassem o problema, dentre eles, os algoritmos variacionais.

É bem sabido que Poincaré já tratara, ainda no século XIX, da possível existência de sistemas dinâmicos extremamente sensíveis às condições iniciais do problema, mas é inegável que a repercussão do trabalho de Lorenz [9] sobre o assunto deu origem a uma área de intensa pesquisa a partir da segunda metade do século XX ([10], [11] e [12]). Considera-se, a seguir, a versão contínua do problema da Assimilação Variacional de Dados 4D.

Os objetos matemáticos disponíveis são:

- as observações do sistema S,  $Z : [t_0, t_1] \rightarrow \mathbb{R}^n$ , sendo a aplicação diferenciável do intervalo de assimilação  $[t_0, t_1]$  em  $\mathbb{R}^n$  munido do produto interno  $\langle \cdot, \cdot \rangle$ .
- a dinâmica do sistema S, dada por

$$\frac{\partial X}{\partial t} = F(X), \quad (2)$$

sendo  $X : [t_0, t_1] \rightarrow \mathbb{R}^n$  a trajetória do sistema S durante o intervalo de tempo  $[t_0, t_1]$ ,  $E = \{Y : [t_0, t_1] \rightarrow \mathbb{R}^n\}$  e  $F : E \rightarrow E$  um operador diferenciável.

- a função "peso"  $W : [t_0, t_1] \rightarrow L(\mathbb{R}^n)$ , sendo  $L(\mathbb{R}^n)$  o espaço vetorial dos operadores lineares em  $\mathbb{R}^n$ , que resulta das informações estatísticas dos instrumentos utilizados na coleta de informações sobre S e que, a cada  $t \in [t_0, t_1]$ , associa o operador linear injetivo  $W(t) : \mathbb{R}^n \rightarrow \mathbb{R}^n$
- o funcional quadrático

$$J : E \rightarrow \mathbb{R} \quad (3)$$

$$X \mapsto \frac{1}{2} \int_{t_0}^{t_1} \langle W(t)(X(t) - Z(t)), W(t)(X(t) - Z(t)) \rangle dt$$

O problema da Assimilação Variacional de Dados 4D é, então, encontrar a trajetória dos estados do sistema S,  $X : [t_0, t_1] \rightarrow \mathbb{R}^n$ , que é isomorfo a  $\{[t_0, t_1] \times \mathbb{R}^n\}$ , tal que X seja a solução de (2) que minimize o funcional (3).

Essencialmente, portanto, X tem de ser tal que, a cada  $t \in [t_0, t_1]$ , se a norma do vetor  $\|W(t)(X(t) - Z(t))\|$  for grande, a norma do vetor erro no instante correspondente,  $\|X(t) - Z(t)\|$ , tem de ser pequena <sup>(1)</sup>, ou seja, se as informações estatísticas de que se dispõe sobre os instrumentos empregados indicam que as observações Z(t) são de alta qualidade, então o erro entre o estado do sistema gerado pelo modelo, X(t), e as observações Z(t) deve ser pequeno. É neste sentido que é dito que os estados X do sistema gerados pelo modelo devem estar próximos das observações Z.

Como se busca a solução numérica da versão discreta do problema anterior, seria computacionalmente muito econômico se, em lugar da solução X, se procurasse a solução  $X_0 \in \mathbb{R}^n$  que minimizasse a restrição do funcional J <sup>(2)</sup> a  $V = \{[t_0, t_1] \times \mathbb{R}^n\} \subset \{[t_0, t_1] \times \mathbb{R}^n\}$ . A dificuldade aqui é não existir uma relação explícita entre J e  $X_0$ , pois J é sim uma função de X. No entanto, se o problema (2) for bem posto, então o conhecimento de X e o de  $X_0$  são equivalentes, e esta é uma informação matemática (e não apenas do particular sistema em estudo) preciosa. Portanto, com todos os ingredientes do problema anterior, quer-se minimizar a restrição do funcional J ao espaço das configurações iniciais do sistema <sup>(2)</sup>. Para tanto, usa-se o método das equações adjuntas.

Restringiu-se, neste parágrafo, aos fundamentos do modelo matemático empregado na Assimilação Variacional de Dados 4D, o que implicou uma simplificação considerável no processo de elaboração de modelos destinados a aplicações ao mundo real. No próximo parágrafo, consideram-se alguns dos aspectos relacionados à implementação de um algoritmo variacional e um exemplo numérico.

### 3 - O MÉTODO DAS EQUAÇÕES ADJUNTAS

Empregar-se-á o método das equações adjuntas para se obter a diferencial da restrição do funcional J ao espaço das configurações iniciais de X, seguindo uma argumentação matematicamente bem definida, desenvolvida pelos autores.

Reescreve-se o funcional (3) como

$$J : E \rightarrow \mathbb{R} \quad (4)$$

$$X \mapsto \int_{t_0}^{t_1} T(X(t)) dt,$$

sendo  $T : E \rightarrow \mathbb{R}$

$$X \mapsto \frac{1}{2} \langle W(t)(X(t) - Z(t)), W(t)(X(t) - Z(t)) \rangle$$

Como a diferencial de J em  $X \in E$  é um funcional linear no espaço de Hilbert E com o produto interno é possível empregar-se um resultado da Análise Funcional, conhecido como o Teorema da Representação de Riesz [13], obtendo-se,  $\forall H \in E$ ,

$$\langle X, Y \rangle_E = \int_{t_0}^{t_1} \langle W(t)(X(t)), W(t)(Y(t)) \rangle dt,$$

$$dJ(X) \cdot H = \langle \nabla_X J(X), H \rangle_E = \int_{t_0}^{t_1} \langle \nabla_X T(X(t)), H(t) \rangle dt, \quad (5)$$

sendo  $\nabla_X J$  e  $\nabla_X T$ , respectivamente, o gradiente de  $J$  e o de  $T$ .

Considera-se, a seguir, a versão linear da equação (2)

$$\frac{\partial X}{\partial t} - \frac{\partial F}{\partial X} X = 0, \quad (6)$$

que resulta da substituição do operador  $F$  pela sua aproximação de primeira ordem em torno de um ponto de equilíbrio, omitido na equação, e na qual se empregou a notação  $(\frac{\partial F}{\partial X})$  para sua diferencial neste ponto, e a seguinte equação adjunta

$$\frac{\partial X}{\partial t} + (\frac{\partial F}{\partial X})^* X + \nabla_X T = 0 \quad (7)$$

sendo  $(\frac{\partial F}{\partial X})^*$  o adjunto do operador  $(\frac{\partial F}{\partial X})$ .

Sejam  $Y : [t_0, t_1] \rightarrow \mathbb{R}^n$  a solução de (7) tal que  $Y(t_1) = 0$  e  $X$  uma solução qualquer de (6). Tem-se, então que

$$\begin{aligned} & \langle \frac{\partial X}{\partial t} - \frac{\partial F}{\partial X} X, Y \rangle + \langle \frac{\partial Y}{\partial t} + (\frac{\partial F}{\partial X})^* Y + \nabla_X T, X \rangle = 0 \\ & \frac{\partial}{\partial t} \langle Y, X \rangle + \langle \nabla_X T, X \rangle = 0 \\ & \therefore \frac{\partial}{\partial t} \langle Y, X \rangle = - \langle \nabla_X T, X \rangle \\ & \langle Y, X \rangle \Big|_{t_0}^{t_1} = - \int_{t_0}^{t_1} \langle \nabla_X T, X \rangle dt = - \langle \nabla_X J, X \rangle \\ & \therefore \langle Y_0, X_0 \rangle = \langle \nabla_X J, X \rangle. \end{aligned} \quad (8)$$

O significado de (6) deve ficar claro: o lado direito de (8) é a expressão da diferencial de  $J$  (em relação a  $X$ ),  $dJ(X)$ . Como  $X_0$  é a projeção sobre  $\mathbb{R}^n$  de  $X$ , o lado esquerdo de (8) é a expressão da diferencial da restrição de  $J$  (em relação a  $X_0$ ). Pela unicidade do vetor gradiente, que permite a representação desta última diferencial como um produto interno em  $\mathbb{R}^n$ , tem-se que  $Y_0$  é precisamente o gradiente da restrição do funcional  $J$  ao  $\mathbb{R}^n$ , ou seja,  $Y_0$  é a projeção de  $\nabla_X J$  sobre o  $\mathbb{R}^n$ , isto é,

$$Y_0 = \nabla_{X_0} J. \quad (9)$$

Uma vez obtido  $\nabla_{X_0} J$ , e com uma configuração inicial  $X_0$  escolhida, é possível buscar-se a perturbação  $H \in \mathbb{R}^n$  de  $X_0$  tal que a trajetória  $X$ , solução de (2) e determinada pela configuração inicial  $(X_0 + H)$ , minimize (3). Para tanto, o conhecimento de  $\nabla_{X_0} J$  será utilizado juntamente com o de  $J(X)$  e de  $X_0$ , numa rotina de minimização baseada em gradientes, como, por exemplo, a L-BFGS [14], empregada em problemas de larga escala, a fim de determinar-se a configuração inicial com as propriedades requeridas. Note-se que o gradiente da restrição do funcional  $J$  ao  $\mathbb{R}^n$ ,  $\nabla_{X_0} J$ , passado à rotina de minimização, nada mais é que a solução da equação adjunta (7) com a condição  $Y(t_1) = 0$ .

Enfatize-se que é a possibilidade de representar-se a diferencial de um funcional linear definido num espaço de Hilbert por meio de um produto interno que permite a realização do cômputo do gradiente da restrição do funcional  $J$  ao  $\mathbb{R}^n$  (ou a uma das cópias do espaço das configurações iniciais de  $S$ , se se preferir) sem o conhecimento explícito da expressão de  $J$  em relação à configuração inicial  $X_0$ . Certamente, este não é um procedimento de menor importância. Os algoritmos variacionais têm possibilitado o tratamento computacional de problemas ambientais cujo número de variáveis é da ordem de  $10^7$  a  $10^9$ . É também utilizado numa gama considerável de problemas, indo desde previsões em finanças ao projeto de estruturas aerodinâmicas ótimas, passando pela dosagem de quimioterápicos necessária ao tratamento de cânceres com um mínimo de efeitos colaterais.

Como aplicação, os autores deste trabalho empregaram o método da Assimilação Variacional de Dados 4D na determinação de uma configuração do sistema formado pelo vazamento acidental de óleo no mar, conhecendo-se um conjunto de observações do sistema resultante (a mancha de óleo formada na superfície do mar) e de sua dinâmica (do meio fluido/óleo), com o objetivo de simular computacionalmente a evolução da mancha de óleo no mar, como um primeiro passo para tomadas de ações visando remediar o problema.

A partir de uma configuração inicial  $X_0$  conhecida, o modelo foi integrado no tempo e, em alguns instantes de tempo  $t_i$  escolhidos no intervalo de assimilação  $[t_0, t_1]$ , as configurações do sistema (mancha de óleo na superfície do mar) foram armazenadas como as observações  $Z_i$  do sistema. Admitiu-se, ainda, a dinâmica do sistema dada pela equação de advecção-difusão 2D:

$$\frac{\partial C}{\partial t} + u \frac{\partial C}{\partial x} + v \frac{\partial C}{\partial y} - D \left( \frac{\partial^2 C}{\partial x^2} + \frac{\partial^2 C}{\partial y^2} \right) = 0, \quad (10)$$

sendo  $\vec{v} = (u, v)$  a velocidade do meio,  $C$  a concentração do poluente e  $D$  o coeficiente de difusividade.

Estava-se, portanto, em condições de aplicar a Assimilação Variacional de Dados 4D para obter a configuração inicial do sistema, supondo-se não mais conhecida a verdadeira condição  $X_0$  (experimento dos gêmeos idênticos, no qual as observações são também geradas pelo modelo numérico [15]).

Um algoritmo para a Assimilação Variacional de Dados 4D, como o exibido na figura 1, foi empregado, sendo  $G(U)$  a integração no tempo do modelo do sistema, a fim de gerar os estados do sistema produzidos pelo modelo,  $J(G(U))$  a avaliação do erro entre os estados do sistema gerados pelo modelo e as correspondentes observações do sistema,  $D(G(U))$  a linearização do modelo  $G(U)$  e  $[D(G(U))]^*$  o adjunto da linearização do modelo  $G(U)$ . Portanto,  $G$  corresponde à integração da discretização (por diferenças finitas ou por elementos finitos) da equação (10), para frente no tempo, com condição inicial  $U$ ,  $D(G(U))$  é a linearização do código de  $G(U)$  (ver observação 2 adiante) também aplicada à condição inicial  $U$  e  $[D(G(U))]^*$  é o código adjunto ao código  $D(G(U))$  (ver também observação 2). Em particular, no experimento realizado pelos autores, como em todo experimento no qual as observações são geradas pelo mesmo modelo numérico (gêmeos idênticos), os operadores  $W(t)$  foram tomados como o operador identidade no  $\mathbb{R}^n$ .

Uma perturbação aleatória  $H$  foi adicionada à configuração inicial  $X_0$ , resultando daí a nova configuração  $X'_0 = X_0 + H$ , que serviu como a primeira tentativa no ajuste entre os estados produzidos pelo modelo (iniciados em  $X'_0$ ) e as observações  $Z_i$ .

O algoritmo variacional, como o apresentado na figura 1, foi utilizado para tentar recuperar a configuração inicial do sistema, partindo-se sempre de uma configuração inicial perturbada  $X_0 + H$ , obtida da rotina de minimização, cuja trajetória de configurações melhor se ajustasse às observações disponíveis  $\{Z_i\}$ .

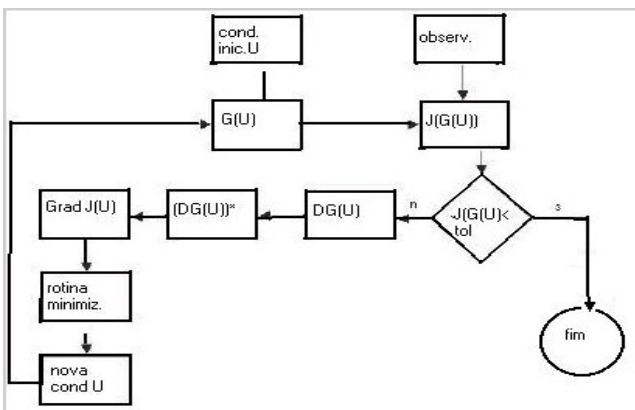


Figura 1: fluxograma para o algoritmo variacional

Nesse caso, o resultado final esperado era que  $H = 0$ , ou seja, que o algoritmo pudesse recuperar a configuração inicial  $X_0$  como a solução do problema, o que de fato ocorreu.

A idéia do experimento era, portanto, responder à pergunta: se são conhecidos um conjunto de observações do sistema e sua dinâmica, o algoritmo variacional é capaz de recuperar sua configuração inicial ou outra qualquer? Como foi visto, a resposta, no experimento de gêmeos idênticos, é sim. Num caso real, portanto, se manchas de óleo são observadas na superfície do mar, onde são conhecidas a velocidade do vento, a hidrodinâmica da região, etc, é possível simular-se a evolução da mancha de óleo e determinar a localização de bóias de contenção num instante de tempo  $t_f > t_1$ .

No caso de vazamento de óleo no mar, o conjunto de instrumentos (satélites, bóias, etc) pode, num certo instante de tempo, capturar apenas parte da mancha resultante, quando se tem, então, apenas uma observação parcial ou incompleta do sistema no instante de tempo considerado.

Uma situação particularmente interessante se dá quando, em cada um dos instantes de tempo no intervalo  $[t_0, t_1]$  em que há observações do sistema disponíveis, estas são incompletas, pois a integração do modelo numérico com uma condição inicial apenas parcial do sistema não simularia a completa evolução do sistema. Este caso foi também analisado pelos autores (figuras 2 a 5) e, desde que o sistema seja completamente observável, a configuração  $X_0$  é recuperada pelo algoritmo variacional (figura 6). Tal situação aproxima-se muito dos problemas reais, uma vez que mesmo com a utilização crescente de satélites, juntamente com outros instrumentos de mensuração já empregados há mais tempo, as observações de uma mancha de poluente derramado acidentalmente no mar são, a cada instante de tempo, incompletas.

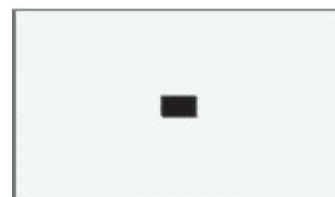


Figura 2: condição inicial do sistema



Figura 3: observação parcial do sistema em  $t = \Delta t \times 50$



Figura 4: observação parcial do sistema em  $t = \Delta t \times 90$



Figura 5: perturbação aleatória na condição inicial

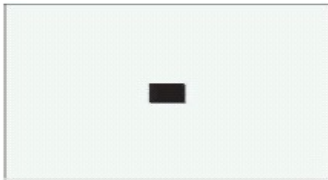


Figura 6: condição inicial recuperada

Algumas observações devem ser feitas:

1. em lugar da equação (2), usamos sua versão linear (6). Este é um procedimento obrigatório se o formalismo adjunto for empregado. Ainda assim, há trabalhos [16] mostrando que, mesmo no caso de sistemas sujeitos a comportamento caótico, o uso de (6) não invalida os resultados obtidos, desde que o intervalo de assimilação seja adequado ao experimento.

2. o formalismo adjunto, por sua vez, criou uma área de intensa pesquisa. De fato, dado um código básico para a resolução de (2), é necessário construir-se o código adjunto que fornece a solução de (5) com  $Y(t_1) = 0$ . A tarefa é, portanto, árdua e exige cautela: se a dinâmica do problema for não linear como em (2), o código para sua solução tem de ser linearizado para daí obter-se o código adjunto. Considere-se um exemplo simples. Seja a seguinte declaração a ser executada no passo de número  $j$  de um dado código em FORTRAN:

$$z = x**2 + \sin(y) \quad (D)$$

Na declaração (D), a variável  $z$  depende não linearmente das variáveis  $x$  e  $y$ , portanto, o primeiro passo é linearizar a declaração.

$dz^{(j)} = 2*x^{(j-1)}*dx^{(j-1)} + \cos(y^{(j-1)})*dy^{(j-1)}$ , (D')

os sobrescritos designando o número do passo no código, que correspondem à informação que estava naquele endereço

de memória no passo em questão. Escrevendo-se (D') em notação matricial, tem-se:

$$\begin{bmatrix} dz \\ dx \\ dy \end{bmatrix}^{(j)} = \begin{bmatrix} 0 & 2*x & \cos(y) \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}^{(j-1)} \begin{bmatrix} dz \\ dx \\ dy \end{bmatrix}^{(j-1)}$$

Como a expressão anterior envolve apenas termos reais, sua adjunta é dada por (3)

$$\begin{bmatrix} dz \\ dx \\ dy \end{bmatrix}^{(j)} = \begin{bmatrix} 0 & 0 & 0 \\ 2*x & 1 & 0 \\ \cos(y) & 0 & 1 \end{bmatrix}^{(j-1)} \begin{bmatrix} dz \\ dx \\ dy \end{bmatrix}^{(j)}$$

que envolve a transposta da expressão anterior.

Assim, a declaração adjunta à (D), (D\*), é dada por

$$\begin{cases} adz^{(j-1)} = 0 \\ adx^{(j-1)} = (2*x^{(j-1)})*adz^{(j)} + adx^{(j)} \\ ady^{(j-1)} = (\cos(y^{(j-1)}))*adz^{(j)} + ady^{(j)} \end{cases}$$

sendo o prefixo *ad* utilizado para designar as variáveis adjuntas *adx*, *ady* e *adz* e evitar confusão entre estas e as variáveis do código básico,  $x$ ,  $y$  e  $z$ . Ou, se abandonam-se os sobrescritos, tem-se de escrever a declaração (D\*) na seguinte ordem:

$$\begin{cases} adx = (2*x)*adz + adx \\ ady = (\cos(y))*adz + ady \\ adz = 0 \end{cases}$$

que é a declaração em FORTRAN adjunta à (D).

Do ponto de vista matemático, talvez todo o processo seja melhor compreendido. De fato, substitua-se a declaração (D) pela função

$$f: \mathbb{R}^3 \rightarrow \mathbb{R}^3 \\ (z, x, y) \mapsto (x^2 + \sin(y), x, y)$$

Então, a primeira expressão matricial anterior é simplesmente a diferencial de  $f$ , dada por sua jacobiana, enquanto que a segunda é a sua transposta. Uma sucessão de declarações seria então tratada com base na aplicação da regra da cadeia à composição de funções correspondente ao conjunto de declarações.

Recorde-se [16] que, dados os espaços vetoriais  $V$  e  $W$  e a transformação linear  $T: V \rightarrow W$  sua transposta é a transformação linear  $T^t: W \rightarrow V$  tal que,  $\forall x \in V$  e  $\forall y \in W$ , tem-se que  $\langle Tx, y \rangle_W = \langle x, T^t y \rangle_V$ . Sendo assim, se pensarmos na ação do código básico à resolução de (2) como uma função que a cada estado inicial  $X_0$  do sistema associa o estado do sistema,  $X_f$ , ao final do intervalo de tempo para a assimilação, então naturalmente seu código adjunto associa a cada estado final do sistema seu correspondente estado inicial, o que justifica tanto a inversão nos índices correspondentes aos passos dos códigos na transposição matricial, quanto a necessidade de impor-se uma condição à solução da equação adjunta (7) no instante final do intervalo de assimilação em vez de uma condição à solução da

equação no instante inicial, como é usual. Estas inversões são características do processo de adjunção.

3. Mencionou-se anteriormente que, de posse de  $\nabla_{X_0} J$  e com uma configuração inicial  $X_0$  escolhida, tem início o processo de Assimilação Variacional de Dados 4D, quando se busca uma perturbação  $H$  de  $X_0$  tal que a trajetória que tem o estado  $(X_0 + H)$  como inicial é a solução ótima do problema no sentido já descrito. O que deve ser ressaltado aqui é que a escolha da configuração inicial  $X_0$  não pode ser completamente aleatória, sob pena de a trajetória resultante não convergir para uma solução, mesmo que o sistema não tenha um comportamento caótico. Se uma das observações  $Z_i$  do sistema for completa, de modo a fornecer uma perfeita configuração do sistema em algum instante de tempo no intervalo  $[t_0, t_1]$ , pode-se tomar  $X_0 = Z_i$ , e iniciar a integração do modelo numérico a partir daí. Porém, quando não há uma observação completa em qualquer dos instantes do intervalo  $[t_0, t_1]$ , é possível, usando-se ainda o formalismo adjunto, encontrar-se uma função de densidade de probabilidade para a configuração inicial do sistema, redefinindo-se adequadamente o funcional  $J$ .

4. A expressão do funcional  $J$ , como definida na seção 2, garante a existência de um mínimo para o problema da Assimilação Variacional de Dados 4D. Ainda assim, em uma sub-região  $R$  da região espacial  $\Delta$  em estudo, mesmo após um número de iterações do algoritmo e com as configurações geradas pelo modelo convergindo em  $\Delta \setminus R$  para as observações disponíveis, pode ocorrer de a discrepância entre estas e aquelas, na região  $R$ , permanecer comparativamente alta ( ver figura 7, onde a região  $R$  está assinalada com um quadrado).

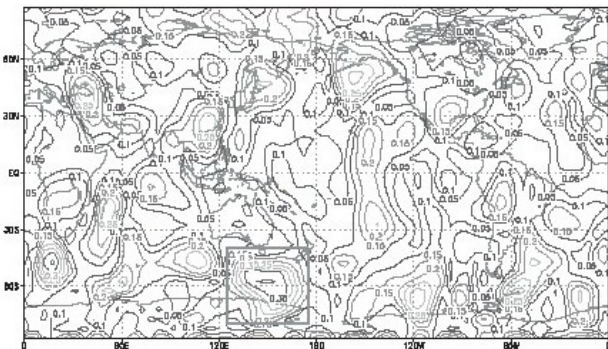


Figura 7: distribuição do erro na norma da energia total no instante de verificação arbitrado

Novamente para entender o que acontece é necessário compreender os fundamentos matemáticos do problema. Ora, se  $G$  representa a integração numérica do modelo, de modo que  $G(X_i) = X_{i+1}$  e  $DG(X_i)$  é a versão linear do modelo no instante  $t_i$ , tal que  $DG(X_i)\delta_i = \delta_{i+1}$ , sendo  $\delta_i$  e

$\delta_{i+1}$  perturbações de  $X_i$  e de  $X_{i+1}$ , respectivamente,  $DG(X_i)$  é um operador linear e seu domínio pode ser decomposto em auto-vetores. Portanto, se a perturbação  $\delta_i$  contiver um erro de observação paralelo a um dos auto-vetores associado a um auto-valor positivo e maior do que 1, este erro será ampliado na perturbação  $\delta_{i+1}$ . Temos assim uma boa indicação para detectar erros nas configurações geradas pelo modelo: identificar os auto-vetores de  $DG(X_i)$  para um instante  $t_i$  adequado e, se necessário, complementar o conjunto de observações com uma distribuição correta de instrumentos segundo a indicação fornecida justamente pelo conhecimento dos auto-vetores de  $DG(X_i)$ , de modo a minimizar os erros de observação responsáveis pelas maiores discrepâncias entre as configurações geradas pelo modelo e as observações do sistema. É aqui temos, mais uma vez, um problema fora do domínio estrito da matemática: a determinação de auto-vetores de matrizes de dimensão muito elevada acarreta um esforço computacional altíssimo, senão inviável. Porém, também este problema pode ser tratado pelo formalismo adjunto a um custo teoricamente muito mais baixo [18], mediante uma definição apropriada do funcional  $J$ .

Um comentário final diz respeito ao fato de o problema descrito na seção 2 independe da natureza das observações. Sendo assim, a Assimilação Variacional de Dados 4D apresenta-se como um método potente à avaliação do desempenho de um modelo computacional na simulação do fenômeno descrito pela equação (2), desempenho que é medido pela discrepância entre os estados produzidos pelo modelo e as observações do sistema em estudo disponíveis, e que pode, nas condições já analisadas, ser ajustado à tolerância estabelecida inicialmente, uma vez haja sido encontrada a configuração inicial  $X_0$  ótima. Ou seja, o desempenho do modelo numérico é tanto melhor quanto mais as configurações por ele produzidas se ajustam às configurações observadas. Sendo assim, os investimentos feitos na produção de projetos eficientes, bem como aqueles na área dos geradores automáticos de códigos lineares e adjuntos (por exemplo, [19] e [20]), justificam-se amplamente.

#### 4 - CONCLUSÕES

O formalismo adjunto, onde o Teorema da Representação de Riesz tem um papel de fundamental importância, contribui notavelmente para a resolução numérica de problemas onde o número de variáveis é muito grande. Procurou-se mostrar como o perfeito entendimento dos fundamentos matemáticos é essencial tanto na identificação de novos problemas quanto em sua resolução; particularmente, na própria derivação dos algoritmos variacionais, na construção de códigos tangentes e adjuntos a um dado código, assim como na detecção de áreas geradoras de discrepâncias entre as configurações geradas pelo modelo e as observações do sistema.

Apresentou-se o formalismo adjunto combinado ao método da Assimilação de Dados 4D como uma eficiente ferramenta tanto na determinação da evolução da trajetória de poluentes lançados em um meio fluido quanto na avaliação do desempenho de modelos numéricos, quando observações do sistema são disponíveis.

Todos os tópicos aqui abordados são ainda objeto de intensa pesquisa de ponta, cujos resultados são determinantes para a solução de problemas de capital importância para a humanidade.

#### REFERÊNCIAS BIBLIOGRÁFICAS

- [1] LIMA, E. L.; Curso de Análise, v.2, Projeto Euclides - IMPA, 1981.
- [2] SORENSEN, H.W.; Parameter estimation: principles and problems, Marcel Dekker, 1980.
- [3] POINCARÉ, H.J.; The value of science: essential writing of H. Poincaré (reunidos por S.J. Gould), Modern Library, 2001.
- [4] L.S. PONTRIAGUIN, V.G. BOLTYANSKI, R.V. GAMKRELIDZE E E.F. MISHCHENKO, The mathematical theory of optimal process, Interscience Publishers, 1971.
- [5] KALMAN, R.E; "A new approach to linear filtering and prediction problems", Transaction of the ASME – Journal of Basic Engineering, pp.: 33-45, 1960.
- [6] LIONS, J.L.; Optimal control of systems governed by partial differential equations, Springer – Verlag, 1971.
- [7] MARCHUK, G.I.; Adjoint equations and analysis of complex systems, Kluwer Academic Publisher, 1994.
- [8] TALAGRAND, O.; "Assimilation of observations, an introduction", Journal of Meteorological Society of Japan, v. 75, n. 1B, pp.: 191-209, 1997.
- [9] LORENZ, E.N.; "Deterministic non-periodic flows", Journal of Atmospheric Sciences, 20, pp.: 130-141, 1963.
- [10] A. KATOK AND B. HASSELBLATT; Introduction to the modern theory of dynamical systems, Cambridge University Press, 1995.
- [11] ROBINSON, C.; Dynamical systems, stability, symbolic dynamics and chaos, CRC Press, 1999.
- [12] MOSER, J.; Stable and random motions in dynamical system with special emphasis on celestial mechanics, Princeton University Press, 1973.
- [13] KREYSZIG, E.; Introductory Functional Analysis with Applications, John Wiley and Sons, 1989.
- [14] D. LIU AND J. NOCEDAL, On the limited memory BFGS method for large scale minimization, Mathematical Programming, 45, pp.: 503-528, 1989.
- [15] O. TALAGRAND; "A study of the dynamics of four-dimensional data assimilation", Tellus, 33, 43-60.
- [16] C. Pires, R. Vautard and O. Talagrand; "On extending the limits of the variational data assimilation in nonlinear chaotic systems", Tellus, 48A, pp.: 96-121, 1996.
- [17] K. HOFFMAN AND R. KUNZE, Álgebra Linear, LTC, 1979.
- [18] D. DAESCU AND I.M. NAVON; "Adaptive observations in the context of 4D-Var data assimilation", Meteorological Atmospheric Physics, 85, pp.: 205-226, 2004.
- [19] R. GIERING AND T. KAMINSKI; "Recipes for adjoint code construction", ACM Transactions on Mathematical Software, v.24, n.4, pp.: 437-474, 1996.
- [20] C. FAURE AND Y. PAPEGAY, Odyssee User's Guide Version 1.7, INRIA, 1998.

#### AGRADECIMENTOS

Um dos autores, P.S.D.S, gostaria de agradecer à gentil atenção do dr. Olivier Talagrand, do Laboratoire de Météorologie Dynamique, França, que, sem qualquer compromisso, jamais deixou uma única de suas muitas perguntas sem resposta.