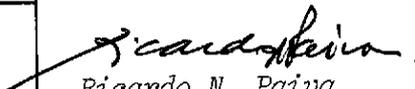
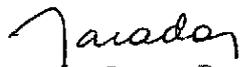


1. Publicação nº <i>INPE-3182-PRE/547</i>	2. Versão	3. Data <i>Julho, 1984</i>	5. Distribuição <input type="checkbox"/> Interna <input checked="" type="checkbox"/> Externa <input type="checkbox"/> Restrita
4. Origem <i>DMC/DDO</i>	Programa <i>ANACO</i>		
6. Palavras chaves - selecionadas pelo(s) autor(es) <i>SUAVIZAÇÃO ESTIMAÇÃO ÓTIMA</i> <i>DOMÍNIO DA FREQUÊNCIA PROCESSOS ESTOCÁSTICOS</i>			
7. C.D.U.: <i>681.5.015.42</i>			
8. Título <i>SUAVIZAÇÃO NO DOMÍNIO DA FREQUÊNCIA</i>		10. Páginas: <i>07</i>	
		11. Última página: <i>06</i>	
		12. Revisada por	
9. Autoria <i>Roberto Vieira da Fonseca Lopes</i>		 <i>Ricardo N. Paiva</i>	
Assinatura responsável 		13. Autorizada por  <i>Nelson de Jesus Parada</i> Diretor Geral	
14. Resumo/Notas <p><i>Diversas contribuições na área de suavização de dados têm possibilitado uma maior utilização desta técnica no processamento em lote de medidas escalares com ruídos. Notadamente, a técnica de suavização por segmentos polinomiais (função "spline"), que permite obter uma solução ótima segundo um dado critério de otimalidade, e de fácil utilização, ainda que, em certas aplicações, requeira um número demasiadamente grande de coeficientes a ser armazenados (proporcional ao número de medidas). O objetivo deste trabalho é obter uma solução subótima para a suavização de dados periódicos, caracterizada por poucos coeficientes em comparação com a solução por "spline". Para tanto, opera-se no domínio da frequência. Também é desenvolvido um suavizador não-periódica para medidas contínuas, o qual se relaciona com o filtro de Butterworth. Uma aplicação com N dados periódicos, uniformemente espaçados, evidencia a economia obtida no número de coeficientes (em termos conservadores, proporcional a \sqrt{N}), sem degradar o desempenho. Os procedimentos desenvolvidos permitem variar a ordem da suavização, que determina a forma da solução, sem alterar a estrutura das equações.</i></p>			
15. Observações <i>Este trabalho será apresentado no 1º Congresso Latino-Americano de Automática/5º Congresso Brasileiro de Automática - 1º CLA/5º CBA - Campina Grande, PB, de 3 a 6 de setembro de 1984.</i>			

SUAVIZAÇÃO NO DOMÍNIO DA FREQUÊNCIA

Roberto Vieira da Fonseca Lopes

Instituto de Pesquisas Espaciais - INPE
Conselho Nacional de Desenvolvimento Científico e Tecnológico - CNPq
C.P. 515 - São José dos Campos - 12.200 - São Paulo - Brasil

Resumo

Diversas contribuições na área de suavização de dados têm possibilitado uma maior utilização desta técnica no processamento em lote de medidas escalares com ruídos. Notadamente, a técnica de suavização por segmentos polinomiais (função "spline"), que permite obter uma solução ótima segundo um dado critério de otimalidade, é de fácil utilização, ainda que, em certas aplicações, requeira um número demasiadamente grande de coeficientes a serem armazenados (proporcional ao número de medidas). O objetivo deste trabalho é obter uma solução subótima para a suavização de dados periódicos, caracterizada por poucos coeficientes em comparação com a solução por "spline". Para tanto, opera-se no domínio da frequência. Também é desenvolvido um suavizador não-periódico para medidas contínuas, o qual se relaciona com o filtro de Butterworth. Uma aplicação com N dados periódicos, uniformemente espaçados, evidencia a economia obtida no número de coeficientes (em termos conservados, proporcional a \sqrt{N}), sem degradar o desempenho. Os procedimentos desenvolvidos permitem variar a ordem da suavização, que determina a forma da solução, sem alterar a estrutura das equações.

Smoothing on the Frequency Domain

Abstract

Several contributions on data smoothing theory have provided a greater utilization of this technique on noisy scalar measurements batch processing. The smoothing technique by spline function, that allows to have an optimal solution in a given sense, has specially easy utilization, nevertheless in some specific cases it demands the storage of an excessive number of coefficients (proportional to the number of measurements). The objective of this work is to construct a suboptimal solution for periodic data smoothing, characterized by few coefficients when compared with spline solution. On this purpose, one works on frequency domain. A nonperiodic smoother for continuous measurements, that is closely related with Butterworth filter is also developed. An application example with N equally spaced data shows the saving in number of coefficients (proportional to \sqrt{N}) obtained without decay of performance. The order of smoothing, that specifies the shape of the solution, can be modified with no modification in the structure of the equations, on both developed procedures.

1. INTRODUÇÃO

Técnicas de suavização de dados com ruídos têm encontrado uma diversidade promissora de aplicações que abrangem problemas em Geofísica, Meteorologia e Medicina, entre outros. Seu uso, além de permitir a redução da incerteza em dados levantados experimentalmente, aplica-se à solução numérica de equações integro-diferenciais e à análise espectral. Uma das técnicas de suavização existentes foi inicialmente formulada por Withaker, em 1923, e vem adquirindo um grau crescente de sofisticação através de vários trabalhos publicados nos últimos anos, como os de Reinsch (1967), De Boor (1978), Wahba (1975), Craven e Wahba (1979) e Utreras Diaz (1980). A base conceitual desta técnica consiste em

dada uma sequência de medidas $\{y(t_i), i = 1, \dots, n\}$ corrompidas por um ruído aleatório, obter a função $g(t)$ que minimiza uma dada norma da diferença entre $g(t)$ e $y(t)$, acrescida de uma dada função penalidade, a qual visa inibir a presença de características espúreas ao fenômeno observado. A solução deve assim satisfazer a um compromisso de equilíbrio entre fidelidade às medidas e às características do fenômeno.

Neste trabalho considera-se o caso, bastante explorado, em que a norma é a média quadrática ponderada pelo inverso do quadrado da incerteza nas medidas, e a função penalidade é:

$$\alpha \int_0^1 [g^{(o)}(t)]^2 dt,$$

onde $\alpha \in \mathbb{R}^+$ e $o \in \mathbb{N}$ são, respectivamente, o grau e a ordem de suavização. A solução deste problema é do tipo "spline" suave polinomial de grau $(2o-1)$ e caracteriza-se por $2o(n-1)$ coeficientes. A principal contribuição deste trabalho é obter uma solução subótima para este mesmo problema, válida para fenômenos periódicos, que se caracteriza por ter poucos coeficientes em comparação com a solução do tipo "spline", sem contudo comprometer a qualidade dos resultados. Para isto restringe-se $g(t)$ a um subespaço conveniente de funções e opera-se no domínio da frequência. Uma outra contribuição refere-se à obtenção de um suavizador não-periódico para o caso de medidas contínuas.

2. DEFINIÇÕES BÁSICAS

Este item dedica-se a definições a ser utilizadas nas Seções 3 e 4.

Seja E_p^o o espaço de funções periódicas do tipo:

$$g(t) = C_0 + \sum_{k=1}^{\infty} [C_k \cos k\omega_0 t + S_k \sen k\omega_0 t], \quad \omega_0 = \frac{2\pi}{T}, \quad (1)$$

onde T é o período e C_k e S_k são reais, tais que:

$$\sum_{k=1}^{\infty} (\omega_0 k)^{2\bar{o}} [C_k^2 + S_k^2], \quad \forall \bar{o} \in [0, o]$$

seja finito. Define-se o seguinte operador:

$$D_p[g(t), o] \triangleq \sum_{k=1}^{\infty} (\omega_0 k)^o [C_k \cos(k\omega_0 t + \frac{\pi}{2} o) + S_k \sen(k\omega_0 t + \frac{\pi}{2} o)], \quad o > 0, \quad g(t) \in E_p^o. \quad (2)$$

Seja E^o o espaço de funções não-periódicas do tipo:

$$g(t) = \frac{1}{\pi} \int_0^{\infty} [C(\omega) \cos \omega t + S(\omega) \sen \omega t] d\omega = \frac{1}{2\pi} \int_{-\infty}^{\infty} G(\omega) e^{-j\omega t} d\omega, \quad (3)$$

onde $G(\omega) = C(\omega) + jS(\omega)$, que é a transformada de Fourier de $g(t)$, é tal que:

$$\int_0^{\infty} \omega^{2\bar{o}} [C^2(\omega) + S^2(\omega)] d\omega, \quad \forall \bar{o} \in [0, o]$$

seja finito. Define-se o seguinte operador:

$$D[g(t), o] \triangleq \frac{1}{\pi} \int_0^{\infty} \omega^o [C(\omega) \cos(\omega t + \frac{\pi}{2} o) + S(\omega) \sen(\omega t + \frac{\pi}{2} o)] d\omega = \frac{1}{2\pi} \int_{-\infty}^{\infty} (-j\omega)^o G(\omega) e^{-j\omega t} d\omega, \quad o \geq 0, \quad g(t) \in E^o. \quad (4)$$

Note-se que os operadores D_p e D apresentam as seguintes propriedades:

$$D_p[ag_1(t) + bg_2(t), o] = aD_p[g_1(t), o] + bD_p[g_2(t), o]$$

$$\frac{d}{dt} D_p[g(t), o-1] = D_p[g(t), o]$$

$$D_p[g(t), o] = \frac{d^o}{dt^o} g(t), \quad \forall o \in \mathbb{N}$$

$$D[ag_1(t) + bg_2(t), o] = aD[g_1(t), o] + bD[g_2(t), o]$$

$$\frac{d}{dt} D[g(t), o-1] = D[g(t), o]$$

$$D[g(t), o] = \frac{d^o}{dt^o} g(t), \quad \forall o \in \mathbb{N}$$

que estabelecem uma analogia com o operador derivada para as funções definidas em E_p^o e E^o , respectivamente, exceto pelo fato de admitirem também ordens $o \in \mathbb{R}$.

Definem-se ainda os seguintes operadores:

$$P_p[g(t), o] \triangleq \frac{1}{T} \int_0^T \{D_p[g(t), o]\}^2 dt, \quad (5) \quad g(t) \in E_p^o,$$

$$P[g(t), o] \triangleq \int_{-\infty}^{\infty} \{D[g(t), o]\}^2 dt, \quad (6) \quad g(t) \in E^o.$$

Substituindo as Definições 2 e 3, respectivamente, nas Definições 5 e 6, obtém-se, após cálculos diretos:

$$P_p[g(t), o] = \frac{1}{2} \sum_{k=1}^{\infty} (k\omega_0)^{2o} [C_k^2 + S_k^2], \quad (7)$$

$$P[g(t), o] = \frac{1}{\pi} \int_0^{\infty} \omega^{2o} [C^2(\omega) + S^2(\omega)] d\omega. \quad (8)$$

Finalizando este item, seja $E_p^{o,K}$ o subespaço de E_p^o definido pela seguinte base:

$$\phi(t) \underline{\Delta} \{1 : \cos \omega_0 t : \dots : \cos K \omega_0 t : \\ : \sin \omega_0 t : \dots : \sin K \omega_0 t\}' \quad (9)$$

composto dos elementos E_p^0 , que possuem C_k e S_k nulos para todo $k > K$. Assim as funções definidas em $E_p^{0,K}$ são do tipo:

$$g(t) = \phi'(t)X,$$

onde $X \in \mathbb{R}^{2K+1}$ é dado por:

$$X = \{1 : C_1 : C_2 : \dots : C_K : \\ : S_1 : S_2 : \dots : S_K\}'$$

3. SUAVIZAÇÃO PERIÓDICA

O problema da suavização periódica pode ser enunciado da seguinte forma: encontre uma aproximação $\tilde{f}_p(t)$ para a função periódica $f(t)$, tal que $\tilde{f}_p(t)$ seja solução de:

$$\text{Inf} \left\{ \sum_{i=1}^n [g(t_i) - y(t_i)]^2 \sigma_i^{-2} + \right. \\ \left. + \alpha P_p[g(t), \sigma] \right\}, \quad g \in E_p^{0,K}, \quad (10)$$

onde $y(t_i)$ é uma medida de $f(t_i)$ corrompida por um ruído aleatório e_i , conforme o seguinte modelo:

$$y(t_i) = f(t_i) + e_i, \quad E\{e_i\} = 0, \\ E\{e_i e_j\} = \begin{cases} \sigma_i^2, & i = j \\ 0, & i \neq j, \end{cases} \quad (11)$$

onde $E\{\cdot\}$ indica o operador esperança.

Utilizando as Equações 1, 7 e 9, pode-se reescrever a Relação 10 na seguinte forma vetorial:

$$\tilde{f}_p(t) = \phi'(t)\tilde{X}, \quad (12)$$

onde \tilde{X} minimiza a seguinte função:

$$J_p(X) = [F'X - Y]' \Xi^{-1} [F'X - Y] + \\ + \beta X' M X, \quad (13)$$

e

$$F \in \mathbb{R}^{n \times 2K+1}, \quad Y \in \mathbb{R}^n, \quad \Xi \in \mathbb{R}^{n \times n},$$

$$M \in \mathbb{R}^{2K+1 \times 2K+1} \text{ e } \beta \text{ são definidos por:}$$

$$F \underline{\Delta} \{\phi(t_1) : \phi(t_2) : \dots : \phi(t_n)\}, \quad (14)$$

$$Y \underline{\Delta} \{y(t_1) : y(t_2) : \dots : y(t_n)\}', \quad (15)$$

$$\Xi \underline{\Delta} \bar{\sigma}_p^2 \cdot \text{diag}(\sigma_1^2 : \sigma_2^2 : \dots : \sigma_n^2), \quad (16)$$

$$\sigma_0^2 \underline{\Delta} n^{-1} \sum_{i=1}^n \sigma_i^2,$$

$$M \underline{\Delta} \text{diag}(0 : 1^{20} : 2^{20} : \dots : K^{20} : \\ : 1^{20} : 2^{20} : \dots : K^{20}), \quad (17)$$

$$\beta \underline{\Delta} \frac{1}{2} \alpha \sigma_0^{-2} \omega_0^{20}. \quad (18)$$

A solução deste problema conduz a um filtro linear em Y , dado por:

$$\tilde{X} = [F \Xi^{-1} F' + \beta M]^{-1} F \Xi^{-1} Y. \quad (19)$$

Utilizando as Equações 12 e 19 e definindo:

$$\tilde{Y} \underline{\Delta} \{\tilde{f}_p(t_1) : \tilde{f}_p(t_2) : \dots : \tilde{f}_p(t_n)\}, \quad (20)$$

$$A \underline{\Delta} F' [F \Xi^{-1} F' + \beta M]^{-1} F \Xi^{-1}, \quad (21)$$

verifica-se que:

$$\tilde{Y} = AY. \quad (22)$$

O fato de \tilde{Y} ser linear em Y permite que o grau de suavização α seja ajustado pelos métodos utilizados, no caso de suavização por "spline", por Craven e Wahba (1979) e Lopes, (1984). Assim sendo, o valor α^* mais indicado para α deve minimizar o estimador não-tendencioso do desvio médio quadrático normalizado de $\tilde{f}_p(t)$ em relação a $f(t)$, $\bar{\sigma}(\alpha)$, definido por:

$$\bar{\sigma}^2(\alpha) \underline{\Delta} \frac{1}{n} \{(\tilde{Y} - Y)' \Xi^{-1} [\tilde{Y} - Y] + \\ + \sigma_0^2 \text{tr}(2A - I)\}, \quad \alpha \in \mathbb{R}^+, \quad (23)$$

caso se conheça o valor de σ_0^2 . Caso contrário, α^* deve, segundo o método de validação cruzada, minimizar a seguinte função:

$$v(\alpha) = [\tilde{Y} - Y]' \Xi^{-1} [\tilde{Y} - Y] \cdot \\ \cdot [\text{tr}(I - A)]^{-2}, \quad \alpha \in \mathbb{R}^+. \quad (24)$$

Em ambos os casos α^* pode ser obtido numericamente por inspeção das Relações 23 ou 24.

Enfocando agora o caso particular em que Ξ é a matriz identidade e os instantes t_i são igualmente espaçados dentro do período, verifica-se que as Equações 19 e 21 reduzem-se a:

$$\tilde{X} = Q^{-1} F Y, \quad (25)$$

$$A = F' Q^{-1} F, \quad (26)$$

onde:

$$\theta(t, \tau) \triangleq \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{e^{j\omega(\tau-t)}}{1 + \alpha(\omega^2)^\sigma} d\omega = \frac{1}{\pi} \int_0^{\infty} \frac{\cos\omega(\tau-t)}{1 + \alpha\omega^{2\sigma}} d\omega \quad (35)$$

Tal como no caso periódico, α pode ser ajustado, a fim de minimizar um estimador $\bar{\sigma}^2(\alpha)$, definido de modo a verificar a seguinte expressão:

$$E\{\bar{\sigma}^2(\alpha)\} = E\left\{ \int_{-\infty}^{\infty} [\bar{f}(t) - f(t)]^2 dt \right\} \quad (36)$$

Utilizando as propriedades estatísticas do ruído em $y(t)$ (descritas pela Equação 30) e a Relação 34, verifica-se que o estimador $\bar{\sigma}^2(\alpha)$ definido por:

$$\bar{\sigma}^2(\alpha) \triangleq \int_{-\infty}^{\infty} [\bar{f}(t) - y(t)]^2 dt - R(0) + \frac{2}{\pi} \int_0^{\infty} \frac{\psi(\omega)}{1 + \alpha\omega^{2\sigma}} d\omega \quad (37)$$

satisfaz à Relação 36, onde $\psi(\omega)$ é a transformada de Fourier de $R(h)$.

5. EXEMPLOS NUMÉRICOS

Até o presente, apenas o caso periódico uniforme foi implementado e testado. Neste item apresenta-se como exemplo ilustrativo do desempenho do procedimento conhecido problema de três corpos - Terra, Lua e a nave "Apolo". O fenômeno observado foi a órbita plana descrita pela nave Apolo no referencial (η, ξ) , cuja origem coincide com o centro de massa do sistema, e o eixo η gira de modo a conter os centros de massa da Terra e da Lua (ver Figura 1), conforme descrito em Shampine (1975). Medidas tomadas uniformemente durante o período T foram obtidas por meio de simulação em computador digital, de acordo com a Equação 11, onde:

$$f(t) = \xi(t)$$

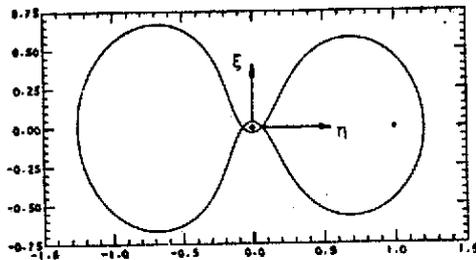


Fig. 1: Órbita da nave "Apolo" no referencial η, ξ .

No caso, utilizou-se $\sigma = 2$. A Figura 2 mostra a trajetória suavizada em comparação

com as medidas no caso de $n = 300$, $K = 30$ e $\sigma_0 = 0,03$.

Para avaliar objetivamente o desempenho atingido, definem-se os parâmetros razão de incertezas, ρ , e número de coeficientes, NC, por:

$$\rho \triangleq \left[\frac{\sigma_0^{-2}}{n} \sum_{i=1}^n [\bar{f}(t_i) - f(t_i)]^2 \right]^{1/2} \quad (38)$$

$$NC \triangleq 2K+1 \quad (39)$$

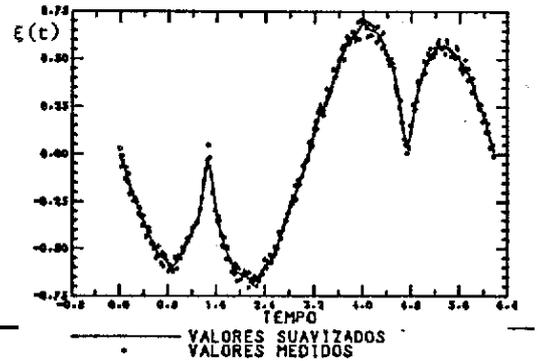


Fig. 2: Exemplo de suavização de 300 medidas com 71 coeficientes.

A Figura 3 mostra a evolução de ρ versus NC para $n = 300$ e $\sigma_0 = 0,03$. As mesmas medidas foram suavizadas pelo método de "spline" polinomial cúbica (1196 coeficientes) utilizando a rotina ICSSCV (IMSL, 1980) e obtendo $\rho = 0,479$, assinalado na Figura 3 por uma linha tracejada. Evidencia-se a vantagem oferecida pelo procedimento em permitir uma redução considerável do número de coeficientes a ser armazenados, os quais caracterizam a curva suavizada, sem comprometer a precisão. Outros testes numéricos realizados indicaram que o número de coeficientes NC*, a partir do qual ρ decresce no máximo 5% de seu valor, cresce com uma potência γ de n tanto menor quanto maior a incerteza nas medidas (ver Figura 4). Em todos os casos testados, NC* cresce sempre com uma potência de N menor que 0,5.

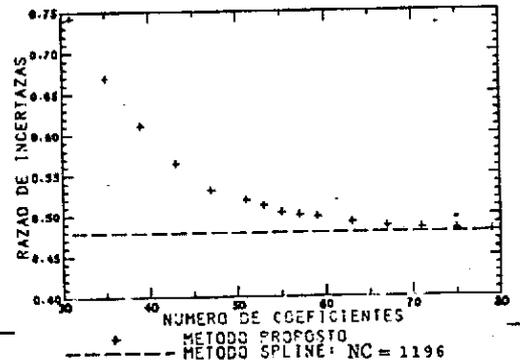


Fig. 3: Desempenho na suavização de 300 medidas: $\rho \times NC$.

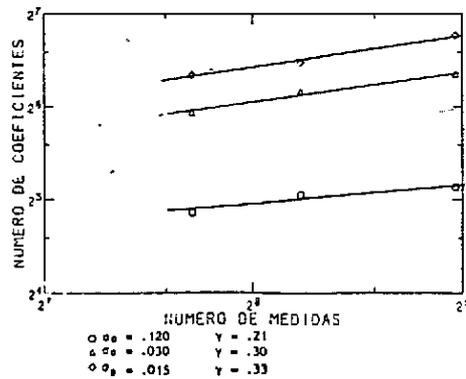


Fig. 4: Variação do número de coeficientes NC^* com o número de medidas N para diferentes valores de incerteza σ_0 .

6. CONCLUSÃO

Desenvolveram-se suavizadores do tipo periódico para medidas sequenciais e não-periódico para medidas contínuas. O caso periódico foi implementado e testado, revelando-se simples e vantajoso para aplicações em que haja interesse em armazenar os resultados, pois para isto torna-se necessário guardar apenas um pequeno número de coeficientes. Os procedimentos são genéricos quanto à ordem de suavização e permitem o ajuste do grau de suavização mais indicado.

REFERÊNCIAS BIBLIOGRÁFICAS

- Craven, P.; Wahba, G. Smoothing noisy data with spline function - estimating the correct degree of smoothing by the method of generalized cross-validation. *Numerische Mathematic*, 31:377-403, 1979.
- De Boor, C. *A practical guide to splines*. New York, Springer Verlag, 1978.
- INTERNATIONAL MATHEMATICAL AND STATISTICAL LIBRARY (IMSL). *Library reference manual*. s.l., 1980, v. 2, Cap. I, p. ICSSCV1 - ICSSCV3.
- Lopes, R.V.F. Suavização de dados: um algoritmo automático. São José dos Campos, INPE, 1984. (INPE-3034-PRE/468). Apresentado com resumo publicado nos anais, nº 69 SBMAC, São José dos Campos, 26 a 30 set., 1983.
- Maybeck, P.S. *Stochastic models, estimation and control*. Vol. 2, Academic Press, New York, 1982.
- Reinsch, C.H. Smoothing by spline functions. *Numerische Mathematic*, 10:177-183, 1967.
- Schwartz, M. *Information transmission, modulation, and noise: an unified approach to communication systems*. Mc Graw-Hill Kogakusha, LTD. Tokyo, 1970, 2ª edição.
- Shampine, L.F.; Gordon, M.K. *Computer solution of ordinary differential equations. The initial value problem*. W.H. Freeman and Company, San Francisco, Chap. 11, 1975.

Utreras Diaz, F. Sur le choix du paramètre d'ajustement dans le lissage par fonctions spline. *Numerische Mathematic*, 34:15-28, 1980.

Wahba, G. Smoothing noisy data with spline functions. *Numerische Mathematic*, 24:383-393, 1975.