

Propriedades da Superfície de Desempenho do Kurtosis Aplicado à Estimação Linear

Pedro I. Hübscher¹ e José C. M. Bermudez^{2*}

¹ INPE - Instituto Nacional de Pesquisas Espaciais
São José dos Campos - SP

² UFSC - Universidade Federal de Santa Catarina, Departamento de Eng. Elétrica
Florianópolis - SC

e-mails: pedro@lit.inpe.br, j.bermudez@ieee.org

Resumo - Este artigo apresenta uma análise das propriedades da superfície de desempenho do problema de estimação linear que utiliza o kurtosis do erro de estimação como função custo. A análise concentra-se na forma modificada da função custo, a qual é empregada na implementação do algoritmo adaptativo Least Mean Kurtosis (LMK). É mostrado como as modificações introduzidas na função custo para viabilizar a implementação prática do algoritmo alteram as propriedades da superfície de desempenho. Essas alterações levam a uma superfície convexa mesmo para sinais de entrada gaussianos. A expressão analítica da superfície de desempenho é determinada como função das propriedades estatísticas dos sinais envolvidos. É demonstrado que essa superfície é unimodal e que o ponto de mínimo corresponde exatamente à solução de Wiener da superfície do erro médio quadrático. Um estudo comparativo entre o erro médio quadrático e a função custo baseada no kurtosis mostra porque o algoritmo LMK converge mais rapidamente do que o algoritmo Least Mean Square (LMS) durante a fase de aprendizado, tornando-se mais lento próximo ao regime permanente.

I. INTRODUÇÃO

Estimação linear tem um papel de grande importância em vários problemas de processamento e controle adaptativos [1], [2], [3]. O projeto otimizado de sistemas adaptativos de estimação requer um conhecimento detalhado do problema teórico e das propriedades do algoritmo adaptativo empregado. Este conhecimento é obtido através da análise do comportamento do sistema, geralmente com a derivação de modelos analíticos que possam prever com boa exatidão o comportamento do algoritmo adaptativo quando aplicado ao sistema.

O estudo do comportamento de um algoritmo adaptativo inclui a previsão da trajetória de seus coeficientes, durante

a fase de aprendizado e em regime permanente, na busca por um ponto estacionário da superfície de desempenho do problema de estimação linear. A eficiência do algoritmo é avaliada, por exemplo, através da comparação do seu desempenho em regime permanente com as propriedades dos pontos estacionários da superfície de desempenho. Assim, o conhecimento das propriedades da superfície de desempenho é necessário para uma avaliação do desempenho de um algoritmo adaptativo, assim como para a comparação dos desempenhos de diferentes algoritmos na solução de um mesmo problema de estimação linear.

A superfície de desempenho mais empregada em problemas de estimação linear é a do erro médio quadrático (EMQ). Em geral, o EMQ apresenta uma dependência de segunda ordem em relação aos coeficientes do filtro adaptativo, tem um único mínimo global, e é matematicamente tratável. Estudos recentes, entretanto, têm mostrado que funções objetivo baseadas em momentos de ordem superior a dois do erro de estimação podem apresentar características que levam a um desempenho superior de algoritmos adaptativos, quando comparado ao desempenho de algoritmos baseados no EMQ. Uma dessas funções objetivo é baseada no *kurtosis* do erro de estimação [2]. O algoritmo adaptativo resultante da utilização dessa superfície de desempenho é denominado *Least Mean Kurtosis*, ou LMK [8], [7].

O algoritmo LMK minimiza uma aproximação para o negativo do *kurtosis* (cumulante de quarta ordem) do erro de estimação. O LMK é um membro da família dos algoritmos do gradiente estocástico [8], [9]. A motivação inicial para a utilização do *kurtosis* veio da propriedade dos cumulantes de ordem superior a dois, os quais são nulos para processos gaussianos de média zero. Como os ruídos de medição em uma grande quantidade de aplicações tendem a processos gaussianos, estimações baseadas no *kurtosis* não são afetadas pelo ruído de medição [11]. No caso de filtragem adaptativa, no entanto, aspectos práticos de implementação forçam a aproximação do *kurtosis*, que pode ser recursiva [8] ou não recursiva [7]. Essas aproximações, ao mesmo tempo que fazem com que a superfície de desempenho desvie da função *kurtosis*, levam a um desempenho

* Este trabalho foi em parte financiado pelo CNPq (Conselho Nacional de Pesquisas), processo número 352084/92-8.

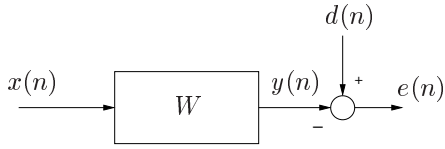


Fig. 1. O problema de estimação linear

do algoritmo LMK que pode superar o de algoritmos baseados no EMQ mesmo para sinais gaussianos [7]. Resultados prévios mostram que o algoritmo LMK pode apresentar desempenho melhor que o algoritmo *Least Mean Square* (LMS) para ruídos de medição com diversas distribuições, inclusive a gaussiana [7], [8], [11]. Além disso, a complexidade computacional dos dois algoritmos é bastante similar para filtros adaptativos de alta ordem. Assim, o estudo do comportamento do algoritmo LMK tem despertado interesse recente na literatura.

A referência [7] apresentou uma análise estatística do comportamento do algoritmo LMK para sinais de entrada gaussianos e para ruídos de medição com funções densidade de probabilidade pares, mas não necessariamente gaussianas. O modelo analítico apresentado em [7] é capaz de prever com bastante exatidão o comportamento do algoritmo, tanto durante a fase de aprendizado quanto em regime permanente. Entretanto, a completa avaliação das propriedades do algoritmo dependem do conhecimento das características da superfície de desempenho utilizada no processo de estimação linear.

Este artigo estuda as propriedades da superfície determinada pela *kurtosis* e da superfície correspondente ao algoritmo LMK realmente implementado. A expressão geral da superfície de desempenho mostra claramente a sua dependência em relação às propriedades estatísticas do sinal de entrada e do ruído de medição. Os pontos estacionários da superfície de desempenho são determinados. A análise do Hessiano dessa superfície mostra que existe apenas um mínimo global, e que esse mínimo corresponde exatamente ao ponto de mínimo da superfície de desempenho do erro médio quadrático. O estudo das propriedades dessa superfície de desempenho permite explicar porque o algoritmo LMK tem uma convergência mais rápida do que o algoritmo LMS no início da adaptação, mas torna-se mais lento próximo ao regime permanente.

II. DEFINIÇÃO DO PROBLEMA

O diagrama da Fig. 1 mostra o problema de estimação linear estudado, sendo:

- $x(n)$: sinal de entrada, estacionário, gaussiano, de média zero e variância σ_x^2 ;
- $d(n)$: sinal desejado;
- $y(n)$: saída do filtro;
- $e(n)$: erro de estimação;
- $W = [w_1, w_2, \dots, w_N]^T$: vetor de coeficientes do filtro linear a ser otimizado;

Na análise que segue, assume-se que o sinal desejado $d(n)$ pode ser modelado como

$$d(n) = W^o{}^T X(n) + z(n) \quad (1)$$

onde $W^o = [w_1^0, w_2^0, \dots, w_N^0]^T$ é um vetor constante, $X(n) = [x(n), x(n-1), \dots, x(n-N+1)]^T$ é o vetor de observações do sinal de entrada e $z(n)$ é um ruído aditivo representando o ruído de medição e os erros de modelagem de $d(n) = W^o{}^T X(n)$. O ruído $z(n)$ é assumido estacionário, branco, não necessariamente gaussiano, mas com função densidade de probabilidade par ($f_z(z) = f_z(-z)$). Assume-se também que $z(n)$ é estatisticamente independente de $x(n)$.

III. SUPERFÍCIE DE DESEMPENHO

Como o erro de estimação será de média zero, o seu *kurtosis* é dado por

$$Cum_4[e(n)] = E[e^4(n)] - 3E^2[e^2(n)] \quad (2)$$

onde Cum_4 significa cumulante de quarta ordem. A função custo J_{LMK} empregada no algoritmo LMK é definida como o negativo do cumulante de quarta ordem [8], [11]. Assim,

$$J_{\text{LMK}} = 3E^2[e^2(n)] - E[e^4(n)] \quad (3)$$

O gradiente de J_{LMK} em relação ao vetor de coeficientes W é dado por

$$\begin{aligned} \nabla J_{\text{LMK}} &= \frac{\partial J_{\text{LMK}}}{\partial W} \\ &= 6E[e^2(n)] \frac{\partial E[e^2(n)]}{\partial W} - \frac{\partial E[e^4(n)]}{\partial W} \end{aligned} \quad (4)$$

Seguindo a metodologia de geração de algoritmos de gradiente estocástico, aproxima-se as derivadas $\frac{\partial E[e^2(n)]}{\partial W}$ e $\frac{\partial E[e^4(n)]}{\partial W}$ em (4) pelas derivadas dos valores instantâneos $e^2(n)$ e $e^4(n)$. Assim, obtém-se uma expressão aproximada para o gradiente da superfície de desempenho na n -ésima iteração:

$$\begin{aligned} \hat{\nabla} J_{\text{LMK}}(n) &\approx \left[6E[e^2(n)] \frac{\partial e^2(n)}{\partial W} - \frac{\partial e^4(n)}{\partial W} \right] \\ &= \left[6E[e^2(n)] 2e(n) \frac{\partial e(n)}{\partial W} - 4e^3(n) \frac{\partial e(n)}{\partial W} \right] \\ &= -4 \left\{ 3E[e^2(n)] e(n) - e^3(n) \right\} X(n) \end{aligned} \quad (5)$$

Na implementação em tempo real da expressão (5), $E[e^2(n)]$ deve ainda ser substituído por uma expressão dependente apenas de valores instantâneos das variáveis. Uma possível solução é estimar $E[e^2(n)]$ usando a equação recursiva [8]

$$E[e^2(n)] = \beta E[e^2(n-1)] + e^2(n), \quad 0 < \beta < 1 \quad (6)$$

Outra possibilidade [7] é substituir (6) por sua solução

$$E[e^2(n)] = \sum_{i=0}^n \beta^i e^2(n-i), \quad E[e^2(0)] = e^2(0) = 0 \quad (7)$$

Uma boa aproximação para β suficientemente pequeno e que leva a uma reduzida complexidade de implementação é obtida utilizando-se os três primeiros termos de (7):

$$E[e^2(n)] \approx e^2(n) + \beta e^2(n-1) + \beta^2 e^2(n-2) \quad (8)$$

Substituindo (8) em (5) resulta finalmente uma aproximação instantânea para o gradiente da superfície de desempenho:

$$\hat{\nabla} \hat{J}_{\text{LMK}}(n) = -8e^3(n)X(n) - 12\beta e^2(n-1)e(n)X(n) - 12\beta^2 e^2(n-2)e(n)X(n) \quad (9)$$

onde a notação \hat{J}_{LMK} é usada para identificar a expressão da superfície de desempenho modificada. A partir de (9) pode-se obter uma expressão para \hat{J}_{LMK} .

Da Fig. 1 tem-se que $e(n) = d(n) - X^T(n)W$. Assim, $\partial e(n)/\partial W = -X(n)$ e (9) pode ser escrita como

$$\begin{aligned} \hat{\nabla} \hat{J}_{\text{LMK}}(n) &= 8e^3(n) \frac{\partial e(n)}{\partial W} \\ &+ 12\beta e^2(n-1)e(n) \frac{\partial e(n)}{\partial W} \\ &+ 12\beta^2 e^2(n-2)e(n) \frac{\partial e(n)}{\partial W} \end{aligned} \quad (10)$$

Integrando (10) e substituindo $e^2(n)$ e $e^4(n)$ por seus valores esperados, obtém-se:

$$\begin{aligned} \hat{J}_{\text{LMK}} &= 2E[e^4(n)] + 6\beta E[e^2(n-1)]E[e^2(n)] \\ &+ 6\beta^2 E[e^2(n-2)]E[e^2(n)] \end{aligned} \quad (11)$$

Para $x(n)$ e $z(n)$ estacionários, $e(n)$ é também estacionário e $E[e^2(n-2)] = E[e^2(n-1)] = E[e^2(n)]$. Definindo o vetor de erro nos coeficientes como $V = W - W^o$ e sendo o erro de estimação dado por $e(n) = z(n) - X^T(n)V$, a expressão da superfície de desempenho pode ser escrita como:

$$\begin{aligned} \hat{J}_{\text{LMK}} &= 2E[e^4(n)] + 6(\beta + \beta^2)E^2[e(n)] \\ &= 2\left\{ E[z^4(n)] + 6\sigma_z^2 V^T R V + 3V^T R V V^T R V \right\} \\ &+ 6(\beta + \beta^2) \left\{ \sigma_z^4 + 2\sigma_z^2 V^T R V + V^T R V V^T R V \right\} \\ &= 6(1 + \beta + \beta^2) \left\{ 2\sigma_z^2 V^T R V + V^T R V V^T R V \right\} \\ &+ 2E[z^4(n)] + 6(\beta + \beta^2)\sigma_z^4 \end{aligned} \quad (12)$$

Neste ponto é importante observar que as aproximações feitas na determinação do gradiente da superfície de desempenho, de forma a viabilizar uma implementação prática

(aproximação estocástica das derivadas e (8)), correspondem a alterações na superfície de desempenho. Em [7] mostra-se que essa aproximação leva a um desempenho do algoritmo LMK que é freqüentemente superior ao do algoritmo LMS. Se o erro médio quadrático $E[e^2(n)]$ fosse substituído pelo seu valor instantâneo (equivalente a fazer $\beta = 0$ em (6)), obter-se-ia a superfície de desempenho do algoritmo *Least Mean Fourth* (LMF) [13], [14]. Finalmente, a aproximação de $E[e^2(n)]$ por (8) torna o comportamento do algoritmo dependente das características estatísticas do sinal de entrada e do ruído aditivo, mesmo se ambos forem gaussianos. Se nenhuma das aproximações fosse feita (esquecendo por um momento a viabilidade de implementação), o algoritmo seria inoperante para sinais de entrada gaussianos. Esses aspectos serão analisados em maiores detalhes na próxima seção, onde são estudadas as propriedades da superfície de desempenho.

IV. PROPRIEDADES DA SUPERFÍCIE DE DESEMPENHO DO *Kurtosis*

Para efeito de comparação, será obtida inicialmente a expressão da superfície de desempenho J_{LMK} em (3) como função das estatísticas do sinal de entrada e do ruído aditivo, especialmente para o caso de sinal de entrada gaussiano.

Sendo o vetor de erro nos coeficientes definido como $V = W - W^o$ e usando $e(n) = z(n) - X^T(n)V$ em (3), leva à expressão para a superfície de desempenho do *kurtosis*:

$$J_{\text{LMK}} = 3E^2[(z(n) - X^T(n)V)^2] - E[(z(n) - X^T(n)V)^4] \quad (13)$$

Em seguida são calculados os valores esperados de (13), considerando a independência estatística entre $z(n)$ e $x(n)$.

Valor Esperado 1:

$$\begin{aligned} E^2[(z(n) - X^T(n)V)^2] &= E^2[z^2(n) - 2z(n)X^T(n)V + X^T(n)VX^T(n)V] \\ &= \left\{ E[z^2(n)] - 2E[z(n)]E[X^T(n)]V + V^T E[X(n)X^T(n)]V \right\}^2 \end{aligned} \quad (14)$$

Definindo $E[z^2(n)] = \sigma_z^2$ e $R = E[X(n)X^T(n)]$ (matriz de autocorrelação de $X(n)$),

$$\begin{aligned} E^2[(z(n) - X^T(n)V)^2] &= \left\{ \sigma_z^2 + V^T R V \right\}^2 \\ &= \sigma_z^4 + 2\sigma_z^2 V^T R V + V^T R V V^T R V \end{aligned} \quad (15)$$

Valor Esperado 2:

$$E[(z(n) - X^T(n)V)^4] = E\left\{ \sum_{i=0}^4 \binom{4}{i} z^i(n) (-X^T(n)V)^{4-i} \right\}$$

Expandindo o somatório, com $E[z^i(n)] = 0$ para i ímpar e considerando $z(n)$ independente de qualquer sinal, tem-

se:

$$E[(z(n) - X^T(n)V)^4] = E[(-X^T(n)V)^4] + 6E[z^2(n)]E[(-X^T(n)V)^2] + E[z^4(n)] \quad (16)$$

A expressão (16) é válida para $x(n)$ com qualquer função densidade de probabilidade. No caso de sinais de entrada gaussianos, $X^T(n)V$ também terá uma distribuição gaussiana e o momento de quarta ordem pode facilmente ser escrito em função dos momentos de segunda ordem. Portanto, para entrada gaussiana, o valor esperado de (16) é dado por

$$\begin{aligned} E[(z(n) - X^T(n)V)^4] &= E[(-X^T(n)V)^4] + 6E[z^2(n)]E[(-X^T(n)V)^2] \\ &+ E[z^4(n)] \\ &= 3E^2[(X^T(n)V)^2] + 6E[z^2(n)]E[(-X^T(n)V)^2] \\ &+ E[z^4(n)] \end{aligned} \quad (17)$$

$$\begin{aligned} &= 3\left\{V^T E[X(n)X^T(n)]V\right\}^2 \\ &+ 6E[z^2(n)]V^T E[X(n)X^T(n)]V + E[z^4(n)] \\ &= 3V^T R V V^T R V + 6\sigma_z^2 V^T R V + E[z^4(n)] \end{aligned}$$

Substituindo-se (15) e (17) em (13), obtém-se:

$$\begin{aligned} J_{\text{LMK}} &= 3\left\{V^T R V V^T R V + 2\sigma_z^2(n)V^T R V + \sigma_z^4\right\} \\ &- \left\{3V^T R V(n)V^T R V + 6\sigma_z^2(n)V^T R V + E[z^4(n)]\right\} \\ &= 3\sigma_z^4 - E[z^4(n)] \end{aligned} \quad (18)$$

A equação (18) corresponde a um valor constante e igual ao negativo do cumulante de quarta ordem (*kurtosis*) do ruído aditivo. Seu gradiente em relação aos coeficientes do filtro adaptativo é identicamente nulo, independentemente do valor desses coeficientes. Neste ponto é importante lembrar que a expressão (18) foi obtida para um sinal de referência gaussiano. Para outras distribuições do sinal de referência, a superfície de desempenho poder ter um formato diferente. Isto não ocorre para a superfície de desempenho do erro médio quadrático, que é a mesma para diferentes funções densidade de probabilidade e sinais de entrada com média zero. Assim, a proposta de utilização do cumulante de quarta ordem como função custo leva a uma superfície de desempenho que não é útil para sinais de entrada gaussianos, mesmo que o ruído aditivo não seja gaussiano.

Na próxima seção são estudadas as propriedades da superfície de desempenho modificada \hat{J}_{LMK} em (12).

V. PROPRIEDADES DA SUPERFÍCIE DE DESEMPENHO MODIFICADA

A. Pontos Estacionários de \hat{J}_{LMK}

Para determinação dos pontos estacionários de (12), deve-se derivar a expressão em relação ao vetor erro dos

coeficientes e igualar o resultado a zero. Os vetores V que satisfazem esta condição podem corresponder a pontos de máximo, mínimo ou pontos de inflexão. A determinação do Hessiano da superfície de desempenho irá indicar em qual destas situações os pontos estacionários se enquadram.

Derivando-se (12) em relação a V e igualando-se o resultado a zero, tem-se:

$$\begin{aligned} \frac{\partial \hat{J}_{\text{LMK}}}{\partial V} &= 6(1 + \beta + \beta^2) \frac{\partial (V^T R V)^2}{\partial V} \\ &+ 12(1 + \beta + \beta^2) \sigma_z^2 \frac{\partial V^T R V}{\partial V} \\ &+ \frac{\partial}{\partial V} \left(2E[z^4(n)] + 6(\beta + \beta^2) \sigma_z^4 \right) \\ &= 0 \end{aligned} \quad (19)$$

A derivadas em (19) são dadas por

$$\begin{cases} \frac{\partial (V^T R V)}{\partial V} = R V + R^T V = 2R V \\ \frac{\partial (V^T R V)^2}{\partial V} = 2(V^T R V) \frac{\partial V^T R V}{\partial V} = 4V^T R V R V \\ \frac{\partial}{\partial V} \left(2E[z^4(n)] + 6(\beta + \beta^2) \sigma_z^4 \right) = 0 \end{cases} \quad (20)$$

Substituindo (20) em (19), tem-se

$$\frac{\partial \hat{J}_{\text{LMK}}}{\partial V} = 24(1 + \beta + \beta^2) \left(\sigma_z^2 + V^T R V \right) R V = 0 \quad (21)$$

Como R é semi-definida positiva [1], a única solução possível para (21) é $V = 0$. Assim, o único ponto estacionário de \hat{J}_{LMK} corresponde a $W = W^o$.

Para determinar se esse ponto estacionário é um ponto de mínimo, máximo ou ponto de inflexão, deve-se determinar o Hessiano de \hat{J}_{LMK} em $V = 0$. Assim,

$$\begin{aligned} H \Big|_{V=0} &= \frac{\partial^2 \hat{J}_{\text{LMK}}}{\partial V^2} \Big|_{V=0} = \frac{\partial}{\partial V} \left(\frac{\partial \hat{J}_{\text{LMK}}}{\partial V} \right) \Big|_{V=0} \\ &= \frac{\partial}{\partial V} \left(24(1 + \beta + \beta^2) \sigma_z^2 R V \right) \Big|_{V=0} \\ &+ \frac{\partial}{\partial V} \left(24(1 + \beta + \beta^2) V^T R V R V \right) \Big|_{V=0} \\ &= 24(1 + \beta + \beta^2) \sigma_z^2 R \Big|_{V=0} \\ &+ 24(1 + \beta + \beta^2) \frac{\partial (V^T R V R V)}{\partial V} \Big|_{V=0} \\ &= 24(1 + \beta + \beta^2) \sigma_z^2 R \end{aligned} \quad (22)$$

O Hessiano será definido positivo se R for definida positiva. Teoricamente, existe apenas a garantia de que R seja semi-definida positiva. Entretanto, na grande maioria das aplicações práticas R será definida positiva devido à inevitável presença de ruído aditivo no sinal $x(n)$ [1]. Portanto, assumindo R definida positiva, conclui-se que H é definida positiva e $V = 0$ é um ponto de mínimo. Substituindo $V = 0$ em (12) tem-se:

$$\hat{J}_{\text{LMK}_{\text{min}}} = 2E[z^4(n)] + 6(\beta + \beta^2) \sigma_z^4 \quad (23)$$

Note que o valor do mínimo dado por (23) é diferente do obtido em (18). isto é consequência das modificações introduzidas na superfície de desempenho J_{LMK} para viabilizar a implementação prática do algoritmo LMK. É também importante notar que o mínimo de \hat{J}_{LMK} corresponde exatamente à solução de Wiener para a minimização do EMQ.

B. Obtenção do Erro Médio Quadrático em Função de \hat{J}_{LMK}

Na comparação de desempenho de diferentes algoritmos adaptativos torna-se necessária a utilização de uma única figura de mérito. Essa figura de mérito é o EMQ na grande maioria das vezes, já que a maioria dos algoritmos adaptativos existentes procura minimizar diretamente alguma função custo proporcional à energia dos erros de estimação. Assim, é de grande interesse a determinação das relações existentes entre as superfícies \hat{J}_{LMK} e do EMQ.

A superfície de desempenho correspondente ao erro médio quadrático é dada [4] por

$$\xi = E[e^2(n)] = \sigma_z^2 + V^T R V \quad (24)$$

Então, a partir de (12) pode-se isolar $V^T R V$ e escrever:

$$\begin{aligned} (V^T R V)^2 + 2\sigma_z^2 (V^T R V) \\ + \frac{2E[z^4(n)] + 6(\beta + \beta^2)\sigma_z^4 - \hat{J}_{\text{LMK}}}{6(1 + \beta + \beta^2)} = 0 \end{aligned} \quad (25)$$

Considerando somente a solução positiva de (25), tem-se

$$V^T R V = -\sigma_z^2 + \frac{1}{2} \sqrt{\frac{4(3\sigma_z^4 - E[z^4(n)]) + 2\hat{J}_{\text{LMK}}}{3(1 + \beta + \beta^2)}} \quad (26)$$

Assim, (26) permite obter o erro médio quadrático em função de \hat{J}_{LMK} . Isto é,

$$\begin{aligned} \xi = f(\hat{J}_{\text{LMK}}) = \sigma_z^2 + V^T R V \\ = \frac{1}{2} \sqrt{\frac{4(3\sigma_z^4 - E[z^4(n)]) + 2\hat{J}_{\text{LMK}}}{3(1 + \beta + \beta^2)}} \end{aligned} \quad (27)$$

Substituindo (23) em (27) tem-se $\xi(\hat{J}_{\text{LMK}_{\text{min}}}) = \sigma_z^2$, concordando com a coincidência dos pontos de mínimo das duas superfícies.

C. Comparação de Taxas de Convergência

Outro aspecto importante da comparação entre as superfícies \hat{J}_{LMK} e ξ é a magnitude de seus gradientes para um dado valor de V . De (27) tem-se

$$4\xi^2 = \frac{12\sigma_z^4 - 4E[z^4(n)] + 2\hat{J}_{\text{LMK}}}{C} \quad (28)$$

sendo $C = 3(1 + \beta + \beta^2) > 3$ para $\beta > 0$.

Derivando ambos os lados de (28) em relação a V e considerando o módulo das derivadas, obtém-se

$$\frac{\left| \frac{\partial \hat{J}_{\text{LMK}}}{\partial V} \right|}{\left| \frac{\partial \xi}{\partial V} \right|} = 4C \xi \quad (29)$$

A análise de (29) mostra que podem ocorrer duas situações extremas:

a) No início da convergência e com ξ grande ($\xi \geq \frac{1}{4C}$),

$$\left| \frac{\partial \hat{J}_{\text{LMK}}}{\partial V} \right| \geq \left| \frac{\partial \xi}{\partial V} \right| \quad (30)$$

Neste caso, o algoritmo LMK converge mais rapidamente que o algoritmo LMS.

b) Perto do regime permanente e com ξ pequeno ($\xi < \frac{1}{4C}$),

$$\left| \frac{\partial \hat{J}_{\text{LMK}}}{\partial V} \right| < \left| \frac{\partial \xi}{\partial V} \right| \quad (31)$$

Nesta situação, o algoritmo LMS converge mais rapidamente.

Esta diferença nas taxas de convergência dos dois algoritmos concorda com uma análise feita a partir de suas equações de atualização do vetor erro dos coeficientes. As equações de atualização de $V(n)$ para os algoritmos LMS [4] e LMK [7] são dadas por

$$V_{\text{LMS}}(n+1) = V_{\text{LMS}}(n) + \mu_{\text{LMS}} e(n) X(n) \quad (32)$$

e

$$\begin{aligned} V_{\text{LMK}}(n+1) = V_{\text{LMK}}(n) \\ + \mu_{\text{LMK}} \{2e^2(n) + 3\beta e^2(n-1) + 3\beta^2 e^2(n-2)\} e(n) X(n) \end{aligned} \quad (33)$$

Assim, o algoritmo LMK pode ser interpretado como sendo o algoritmo LMS com o passo de adaptação variante no tempo:

$$\mu_{\text{LMS}}(n) = \mu_{\text{LMK}} \{2e^2(n) + 3\beta e^2(n-1) + 3\beta^2 e^2(n-2)\} \quad (34)$$

Sabe-se que a taxa de convergência e o desajuste do algoritmo LMS são direta e inversamente proporcionais ao passo de adaptação, respectivamente [1], [2]. Idealmente, um algoritmo adaptativo deve ter uma alta taxa de convergência durante a fase transitória e um pequeno desajuste em regime permanente. Portanto, é desejável que o passo de adaptação seja grande no início do processo de adaptação ($e(n)$ grande) e pequeno próximo à convergência ($e(n)$ pequeno). A expressão (34) apresenta esse comportamento.

VI. CONCLUSÕES

Este artigo apresentou uma análise das propriedades da superfície de desempenho do problema de estimação linear que utiliza o *kurtosis* do erro de estimação como função custo. A análise concentrou-se na forma modificada da função custo que é empregada na implementação do algoritmo adaptativo *Least Mean Kurtosis* (LMK). Foi demonstrado que a superfície resultante é unimodal e que o seu mínimo corresponde exatamente à solução de Wiener da superfície do erro médio quadrático. Um estudo comparativo entre o erro médio quadrático e a função custo baseada no *kurtosis* mostrou porque o algoritmo LMK converge mais rapidamente do que o algoritmo LMS durante a fase de aprendizado, tornando-se todavia mais lento próximo ao regime permanente.

REFERÊNCIAS

- [1] S. Haykin, *Adaptive Filter Theory*, 4ª Ed., Upper Saddle River, NJ: Prentice-Hall, 2002.
- [2] D. G. Manolakis, V. K. Ingle e S. M. Kogon, *Statistical and Signal Processing: Spectral Estimation, Signal Modeling, Adaptive Filtering and Array Processing*, Boston, Mass: Mc-Graw Hill, 2000.
- [3] T. Kailath, A. H. Sayed and B. Hassibi, *Linear Estimation*, Upper Saddle River, NJ: Prentice-Hall, 2000.
- [4] B. Widrow and S. Stearns, *Adaptive Signal Processing*, Englewood Cliffs, NJ:Prentice-Hall, 1985.
- [5] D. R. Morgan and S.M. Kuo, *Active Noise Control Systems: Algorithms and DSP Implementations*, first edition, NY:John Wiley and Sons, 1996.
- [6] S. E. Woodard and A. Nagchaudhuri, "Application of least mean square algorithms to spacecraft vibration compensation", *The Journal of the Astronautical Sciences*, v. 48, n.1, p.83-90, Jan-Mar, 1998.
- [7] P. I. Hbscher and J. C. M. Bermudez, "A model for the behavior of the least mean kurtosis (LMK) adaptive algorithm with gaussian inputs", *International Telecommunications Symposium - ITS2002*, Natal/RN, Brasil, 2002.
- [8] O. Tanrikulu and A. G. Constantinides, "Least Mean Kurtosis: a novel higher-order statistics based adaptive filtering algorithm", *Electronic Letters*, 30(3): 189-190, Feb., 1994.
- [9] A. G. Constantinides and D. I. Pazaitis, "A novel kurtosis driven variable step-Case adaptive algorithm", *IEEE Transactions on Signal Processing*, v. 47, n.3, Mar., 1999.
- [10] A. Papoulis, *Probability, Random Variables and Stochastic Processes*, 3. ed., McGraw-Hill, 1991.
- [11] O. Tanrikulu, *Adaptive Signaling Processing Algorithms with Accelerated Convergence and Noise Immunity*, PhD Thesis, Department of Electrical Engineering, Imperial College, London, 1995.
- [12] O. Tanrikulu and A. G. Constantinides, "The LMK algorithm with time-varying forgetting factor for adaptive system identification in additive output-noise", *Proceedings of the ICASSP'96*.
- [13] E. Walach and B. Widrow, "The least mean fourth (LMF) adaptive algorithm and its family", *IEEE Transactions on Information Theory*, 30(2), pp. 275-283, 1984.
- [14] P. I. Hbscher and J. C. M. Bermudez, "An improved stochastic model for the least mean fourth (LMF) adaptive algorithm", *IEEE International Symposium on Circuits and Systems, ISCAS-2002*, Scottsdale, Arizona, USA, 26-27/May, 2002.
- [15] D. Zwillinger, *Standard Mathematical Tables and Formulae*, 30 ed. CRC.