

REDE NEURAL COM FUNÇÃO DE BASE RADIAL COM TREINAMENTO USANDO FILTRO DE KALMAN E ENTROPIA CRUZADA APLICADA À IDENTIFICAÇÃO DE SISTEMAS

LEANDRO DOS SANTOS COELHO

*Laboratório de Automação e Sistemas, Grupo Produtônica
Pós-Graduação em Engenharia de Produção e Sistemas, PUCPR / PPGEPS
Imaculada Conceição, 1155, CEP 80215-901 Curitiba, PR
E-mails: leandro.coelho@pucpr.br*

Abstract Radial basis function neural networks (RBF-NNs) are function approximation models that can be trained by examples to implement a desired input-output mapping (interpolation). The performance of RBF-NNs depends on the number and centers of the radial basis functions, their shapes, and the method used for learning the input-output mapping. This paper presents a RBF-NN model using Gaussian functions and Kalman filter training based on a cross entropy method. The simulation results indicate the potentialities of this RBF-NN learning approach to one step ahead identification of heat flow density (output of thermal system) through a two layer wall (brick and insulation layer). The system inputs are the internal and external temperature of the wall.

Keywords—Radial basis function neural network, systems identification, cross entropy.

Resumo As redes neurais com função de base radial (RNs-RBF) são modelos de aproximação de funções que podem ser treinados por exemplos para implementar um mapeamento (interpolação) de entrada e saída desejado. O desempenho das RNs-RBF depende do número e centros das funções radiais de base, suas formas, e o método utilizado para aprender o mapeamento de entrada-saída. Este artigo apresenta uma abordagem de RN-RBF usando funções Gaussianas e treinamento com filtro de Kalman baseado em otimização usando um método de entropia cruzada. Os resultados de simulação indicam as potencialidades desta abordagem de aprendizado de RN-RBF para identificação de curto prazo (um passo à frente) da densidade de fluxo de calor (saída do sistema térmico) através de duas camadas de parede (tijolo e camada de isolamento). As entradas do sistema térmico são: a temperatura interna e a temperatura externa da parede.

Palavras-chave Rede neural com função de base radial, identificação de sistemas, entropia cruzada.

1 Introdução

A identificação de sistemas é o estudo de procedimentos que permitem desenvolver modelos matemáticos a partir de dados e sinais observados. As redes neurais artificiais são eficientes estruturas não-lineares para identificação do tipo caixa-preta em muitas aplicações (Haykin, 1996). A principal potencialidade na estrutura das redes neurais consiste na sua habilidade de adaptação e aprendizagem. A habilidade de adaptação e aprendizagem pelo ambiente significa que modelos de redes neurais podem lidar com dados imprecisos e situações complexas. Uma rede neural treinada adequadamente tem a habilidade de generalizar quando são apresentadas as entradas que não estão presentes em dados já conhecidos por ela.

Um tipo de rede neural que pode ser aplicado à identificação de sistemas é a rede neural de base radial (RN-RBF, *Radial basis function neural network*). Parte do interesse pelas RNs-RBF é proveniente do fato delas possuírem fundamentação estatística e serem consideradas aproximadores universais de funções com dada precisão (Chen *et al.*, 1991), realizando um mapeamento das entradas de rede de dimensão vetorial m , em uma saída de dimensão vetorial n de qualquer função vetorial contínua.

As RNs-RBF apresentam potencialidades de generalização local, em que a interferência no aprendizado é minimizada e o aprendizado é relativamente rápido, devido ao número mínimo de pesos ajustados a cada vetor de treinamento apresentado à rede neural. Na literatura, diferentes abordagens de treinamento de redes neurais usando filtro de Kalman têm sido propostas, tais como Puskorius e Feldkamp (1994), Birgmeier (1995), Simon (2002), Xiong e Zhou (2003) e Zhang e Luh (2005).

A contribuição deste trabalho é avaliar as potencialidades de usar uma abordagem de treinamento de RNs-RBF baseada em filtro de Kalman estendido (EKF, *Extended Kalman Filter*) com otimização usando uma abordagem de entropia cruzada (Rubinstein, 1997; Rubinstein, 1999) para identificação do comportamento dinâmico de um sistema térmico. Neste contexto, o estudo de caso testado tem por objetivo a estimativa da densidade de fluxo de calor através de uma parede de duas camadas (camada de tijolos e isolamento).

O restante do artigo é organizado da seguinte forma. A descrição e a abordagem de treinamento da RN-RBF usando otimização baseada em entropia cruzada é detalhada na seção 2. Uma breve descrição do sistema térmico e a análise dos resultados obtidos são detalhadas na seção 3. Finalizando, a conclusão é apresentada na seção 4.

2 Rede Neural de Base Radial

A RN-RBF pode ser considerada como um projeto de rede neural para um problema de ajuste de curva (aproximação) em um espaço de alta dimensionalidade.

Para uma RN-RBF, as unidades ocultas fornecem um conjunto de “funções” que consistem em uma “base” arbitrária para os padrões (vetores) de entrada, quando eles são expandidos sobre o espaço oculto: estas funções são denominadas de funções de base radial, das quais a rede deriva seu nome. Cada uma dessas funções é centrada em uma particular coordenada do espaço multidimensional dos pontos que compõem o espaço de dados de entrada. Cada uma destas coordenadas particulares caracteriza-se por definir o centro de uma (entre várias possíveis) região de maior aglomeração de pontos ou grupos (*clusters*), do espaço de dados de entrada (Haykin, 1996).

A arquitetura da RN-RBF, em sua estrutura mais básica, é constituída de três camadas, com propósitos totalmente diferentes:

- *Camada de entrada*: é composta por nós de fonte (unidades sensoriais) que conectam a rede ao seu ambiente;
- *Camada oculta*: é única, aplica uma transformação não-linear do espaço de entrada para o espaço oculto (na maioria das aplicações, o espaço oculto é de alta dimensionalidade);
- *Camada de saída*: é linear, fornecendo a resposta da rede ao padrão (sinal) de ativação aplicada à camada de entrada.

O processamento da entrada de uma rede neural se realiza de acordo com sua topologia e o tipo da ativação dos neurônios. Em uma RN-RBF clássica, este aspecto consiste na alimentação adiante (*feedforward network*) de uma entrada de dimensão m em uma camada oculta, que possui c neurônios ocultos que implementam uma função de ativação baseada em uma medida de similaridade entre os vetores de entrada e os protótipos. Cada neurônio da camada oculta possui um protótipo. As saídas da camada oculta são interligadas às unidades da camada de saída por ramos com pesos sinápticos que ponderam cada saída dos neurônios presentes na camada oculta. Cada unidade da camada de saída implementa uma soma ponderada das saídas da camada oculta. A resposta das camadas de saída forma os elementos do vetor de resposta da RN-RBF.

Uma função de base radial, f_c , apresenta uma saída simétrica em torno de um centro associado, \mathbf{m}_c , tal que $f_c = f(\|x - \mathbf{m}_c\|; \mathbf{s})$, onde $\|\cdot\|$ é a norma vetorial. Uma condição suficiente para uma função ser candidata à função de base radial é a de ser radialmente simétrica e possuir primeira derivada monotônica. A função Gaussiana é o tipo de função de base radial utilizada neste artigo. As funções

Gaussianas são caracterizadas por um parâmetro de escala (ou comprimento), \mathbf{s} , e isto é verdade para muitas outras classes de funções de base radial. Tais classes podem ser representadas como:

$$f_c(x) = f(\|x - \mathbf{m}_c\|; \mathbf{s}) \quad (1)$$

Um conjunto de funções de base radial pode servir como uma base para representação de uma diversificada classe de funções, que são expressas como combinações lineares de funções de base radial, tal que:

$$y(x) = \sum_{j=1}^M w_j f(\|x - \mathbf{m}_c\|; \mathbf{s}) \quad (2)$$

onde w_j são os pesos sinápticos e M é o número de protótipos. Os vetores de protótipos e pesos sinápticos são os parâmetros da estrutura da RN-RBF. Deve-se ressaltar que os valores desses parâmetros não acrescentam nada de conhecimento a respeito do sistema, porém devem cumprir o determinado: aproximar a resposta do sistema a uma determinada entrada. Para tanto a estimação dos parâmetros das redes neurais em geral, incluindo as RNs-RBF pode ser concebido como um problema de aprendizado com uma “possível” interpretação biológica e também com equivalência matemática (problema de otimização não-linear na estimação dos parâmetros do modelo segundo um critério de otimalidade) (Haykin, 1996).

Alguns pesquisadores têm treinado as RNs-RBF, escolhendo os centros das unidades ocultas aleatoriamente no conjunto de treinamento. Outros têm utilizado métodos não-supervisionados (por exemplo, agrupamento pelo método de k -médias e fuzzy c -médias), como também algoritmos supervisionados para a seleção dos centros (Simon, 2002). Os métodos de treinamento separam as tarefas de determinação dos protótipos e a otimização dos pesos. A vantagem dessa abordagem é a rapidez no treinamento da rede como resultado da separação das tarefas e aplicação de um método híbrido de treinamento em alternativa ao treinamento global e a otimização não-linear explícita da descida de encosta (incluindo o algoritmo de retropropagação do erro) (Chen *et al.*, 1991).

A separação das tarefas no treinamento da RN-RBF é possível devido à estrutura da RN-RBF e seu tipo de função de ativação que age localmente no espaço de entrada. Foi dada uma considerável atenção na literatura das RNs-RBF em como obter um método de treinamento baseado em técnicas lineares explorando características estruturais das RNs-RBF.

A estrutura da RN-RBF possibilita a separação de tarefas no processo de aprendizado, com a vantagem que na estimação dos pesos sinápticos w o treinamento da RN-RBF se torna uma tarefa rápida de ser resolvida através de métodos do tipo mínimos quadrados. Neste trabalho, o algoritmo filtro de

Kalman é utilizado para o treinamento da RN-BRF sendo este descrito na subseção a seguir.

2.1. Fundamentos do filtro de Kalman estendido

O filtro de Kalman, proposto por Kalman (1960), foi desenvolvido inicialmente como uma solução recursiva para filtragem linear de dados discretos. Para isto, utiliza equações matemáticas que implementam um estimador preditivo de estados, buscando corrigir iterativamente a resposta de um determinado sistema através de múltiplas variáveis relacionadas a ele. Suas áreas de aplicação são muito diversificadas, tais como: processamento de imagem, supervisores de eventos discretos, processamento de sinais, sistemas de inferência, entre outros (Côrrea, 2005).

O filtro de Kalman original (Kalman, 1960) foi concebido para ser aplicado em sistemas com dinâmica linear. Para um sistema dinâmico linear com ruído branco de processo de medição, o filtro de Kalman é um estimador ótimo dos estados do sistema. A fim de tornar possível a aplicação do filtro de Kalman em sistemas não-lineares, pesquisadores propuseram o filtro de Kalman estendido (Kuchner, 1967), o que se mostrou eficiente para muitos sistemas cuja não-linearidade é pouco significativa. Para sistemas não-lineares com ruído colorido, o filtro de Kalman pode ser estendido pela linearização da matriz de transição não-linear da descrição em variáveis de estado do sistema em torno dos parâmetros estimados correntes (Haykin, 2001). Neste contexto, o filtro de Kalman estendido atualiza os parâmetros de forma a ser consistente com todos os dados medidos do processo e geralmente converge em poucas interações (Simon, 2002).

A formulação dos filtros de Kalman está fundamentada na representação em espaço de estados de um sistema dinâmico linear discreto de dimensão finita, ou seja, o filtro de Kalman é descrito em termos de equações a diferenças, porém se o sistema for não-linear, pode-se estender o uso do filtro de Kalman com um procedimento de linearização de um sistema da forma

$$\begin{aligned} \mathbf{q}_{k+1} &= f(\mathbf{q}_k) + \mathbf{w}_k \\ y_k &= h(\mathbf{q}_k) + \mathbf{u}_k \end{aligned} \quad (3)$$

onde o vetor \mathbf{q}_k é o estado do sistema no instante de tempo K , \mathbf{w}_k é o ruído provocado pela própria dinâmica do processo, y_k é o vetor de observação, \mathbf{u}_k é o ruído de observação (ou de medida), e $f(\cdot)$ e $h(\cdot)$ são funções não-lineares.

A idéia básica do filtro de Kalman estendido consiste em linearizar a equação (3) a cada instante de tempo em torno do estado estimado recente. Se as não-linearidades são suficientemente suaves, pode-se fazê-lo aplicando a aproximação de Taylor de primeira ordem nas funções não-lineares v e $h(\cdot)$ em torno do estado estimado $\hat{\mathbf{q}}$ para obter

$$f(\mathbf{q}_k) = f(\hat{\mathbf{q}}_k) + F_k(\mathbf{q}_k - \hat{\mathbf{q}}_k) + \text{termos de ordem elevada} \quad (4)$$

$$h(\mathbf{q}_k) = h(\hat{\mathbf{q}}_k) + H_k^T(\mathbf{q}_k - \hat{\mathbf{q}}_k) + \text{termos de ordem elevada} \quad (5)$$

sendo adotada a seguinte notação para as derivadas parciais:

$$F_k = \left. \frac{\partial f(\mathbf{q})}{\partial \mathbf{q}} \right|_{\mathbf{q}=\hat{\mathbf{q}}_k} \quad (6)$$

$$H_k^T = \left. \frac{\partial h(\mathbf{q})}{\partial \mathbf{q}} \right|_{\mathbf{q}=\hat{\mathbf{q}}_k} \quad (7)$$

Desconsiderando os termos de ordem mais elevada na equação (5), um sistema linear aproximado é obtido de (4), onde

$$\mathbf{q}_{k+1} = F_k \mathbf{q}_k + \mathbf{w}_k + \mathbf{f}_k \quad (8)$$

$$y_k = H_k^T \mathbf{q}_k + v_k + \mathbf{j}_k \quad (9)$$

Neste caso, \mathbf{f}_k e \mathbf{j}_k são definidos por

$$\mathbf{f}_k = f(\hat{\mathbf{q}}_k) - F_k \hat{\mathbf{q}}_k \quad (10)$$

$$\mathbf{j}_k = h(\hat{\mathbf{q}}_k) - H_k^T \hat{\mathbf{q}}_k \quad (11)$$

A obtenção da aproximação linear do sistema representado pela equação (4) é o primeiro passo para utilização do filtro de Kalman. O problema do filtro de Kalman consiste em resolver simultaneamente e de maneira ótima as equações de processo e de medida para um estado desconhecido.

Pode-se formular o problema de filtragem de um sistema linear achando um valor ótimo de $\hat{\mathbf{q}}_k$, tal que o sistema deve satisfazer as restrições impostas pelos teoremas do estimador médio condicional e o do princípio da ortogonalidade. Finalmente, pode-se mostrar que a estimação desejada $\hat{\mathbf{q}}_k$ pode ser obtida recursivamente pelas equações do filtro de Kalman:

$$\hat{\mathbf{q}}_k = f(\hat{\mathbf{q}}_{k-1}) + K_k [y_k - h(\hat{\mathbf{q}}_{k-1})] \quad (12)$$

$$K_k = P_k H_k (R + H_k^T P_k H_k)^{-1} \quad (13)$$

$$P_{k+1} = F_k (P_k - K_k H_k^T P_k) F_k^T + Q \quad (14)$$

onde K_k é o ganho de Kalman e P_k a matriz de covariância do erro de estimação do estado \mathbf{q}_k . As matrizes Q e R são as matrizes de covariância dos ruídos artificiais \mathbf{w}_k e \mathbf{u}_k , respectivamente. Como o sistema descrito nas equações (8) e (9) é uma aproximação de um sistema não-linear, o filtro não é ótimo e a estimação é somente uma aproximação da média condicional.

O problema de otimização da matriz de pesos sinápticos W e as coordenadas dos vetores protótipos v_j sob o ponto de vista da estimação ótima de parâmetros com filtro de Kalman de uma RN-RBF pode ser tratado como um problema de minimização do tipo mínimos quadrados ponderados. Neste caso, para aplicar o EKF, faz-se com que os elementos da matriz de pesos W e os elementos dos vetores de

protótipos v_j constituam o estado do sistema discreto não-linear em que o filtro de Kalman é aplicado.

2.2. Otimização usando entropia cruzada

Um algoritmo de amostragem de dados adaptativo foi proposto por Rubinstein (1997) para estimar “raros” eventos pela minimização da variância de amostragem do estimador. Mais tarde, este algoritmo foi modificado por Lieber e Rubinstein (1998) para minimizar a distância (divergência) de Kullback-Leibler (ou entropia cruzada), uma bem conhecida medida de informação. Algumas modificações no método de entropia cruzada têm sido propostas na literatura para resolução de problemas de otimização combinatória (Rubinstein, 1997) e problemas de otimização contínua (Rubinsten, 1999). Neste contexto, o método de entropia cruzada pode ser descrito como um algoritmo iterativo que consiste de duas etapas principais a cada iteração: (i) gerar uma amostra de dados aleatórios regidos por um conjunto de parâmetros e uma função densidade de probabilidade, e (ii) atualizar o conjunto de parâmetros que controla a geração de dados aleatórios baseado na qualidade da amostra gerada, visando melhorar os dados na iteração seguinte.

Neste trabalho, utiliza-se um método de entropia cruzada para otimização contínua, ou seja, otimizar as variáveis da diagonal principal de P , Q e R do filtro de Kalman (uma solução x de cada amostra de dados na entropia cruzada) visando o treinamento de uma RN-RBF. As etapas do método de entropia cruzada utilizado podem ser resumidas em:

Etapa 1. Iniciar os parâmetros de controle.

- Contador de iterações: $k = 1$;
- Valor inicial para a média, $\mu_{ec}(k)$:

$$\mu_{ec}(k) = LI - 0,5(LS - LI), \quad (15)$$
 onde LS e LI são os limites inferior e superior das variáveis (soluções) a serem otimizadas;
- Valor inicial para o desvio padrão, $s_{ec}(k)$:

$$s_{ec}(k) = 0,035 \max(\mu_{ec}(k) - LI, LS - \mu_{ec}(k)) \quad (16)$$
- Parâmetro de raridade: $r = 0,2$;
- Tamanho da amostra: $N = 10$;
- Parâmetro de suavizamento da distribuição de probabilidade: $a = 0,7$;
- Parâmetros para atualização do desvio padrão: $b = 0,995$ e $q = 5$.

Etapa 2. Utilizar uma função densidade de probabilidade para gerar uma amostra de dados de tamanho N . Neste trabalho adota-se uma distribuição de probabilidade normal Gaussiana truncada no intervalo $[LI, LS]$ para a média $\mu_{ec}(k)$ e o desvio padrão $s_{ec}(k)$.

Etapa 3. Avaliar a função objetivo de cada solução X que compõe a amostra de dados $\{x_{i1}, \dots, x_{iN}\}$, onde $i=1 \dots D$, onde D é a dimensão do problema (neste trabalho, $D=3$).

Etapa 4. Ordenar a função custo de cada solução da amostra e determinar o quartil da amostra (medida de variabilidade) definido por $(1 - r)$. Isto significa (para $r = 0,2$) uma elite de 20% dos melhores valores de função custo.

Etapa 5. Incrementar o contador de iteração, tal que $k=k+1$;

Etapa 6. Atualizar a média e desvio padrão usando as melhores $r\%$ soluções da amostra X . Neste caso, as seguintes equações foram utilizadas:

Equação para o cálculo de um suavizamento da média:

$$\mu_{ec}(k+1) = a \cdot \mu_{ec}(k) + (1-a) \cdot \mu_{ec}(k) \quad (17)$$

Equações para o cálculo de um suavizamento dinâmico do desvio padrão:

$$B(k+1) = b - b \cdot (1-(1/k))^q \quad (18)$$

$$s_{ec}(k+1) = B(k+1) \cdot s_{ec}(k) + (1-B(k+1)) \cdot s_{ec}(k) \quad (19)$$

Etapa 7. Ir para a *Etapa 2* até que um critério de parada seja atendido. Neste trabalho, adotou-se como critério de parada $k = 30$ iterações.

Etapa 8. Melhor solução obtida é dada pela solução que obteve menor valor de função objetivo (problema de minimização). Neste trabalho, a melhor solução x_{ij} (P , Q e R do filtro de Kalman), onde $i=1, \dots, D$, para $j \in [1, \dots, N]$ obtida pelo método de entropia cruzada será aquela que obtiver o menor erro médio quadrático de previsão da RN-RBF para previsões de um passo à frente.

3 Descrição do Sistema Térmico e Análise dos Resultados de Identificação usando uma RN-RBF

A seguir são apresentados os detalhes do sistema térmico a ser identificado e após são mostrados e comentados os resultados de identificação usando RN-RBF com filtro de Kalman e entropia cruzada.

3.1. Descrição do sistema térmico

A obtenção de uma estimativa apropriada para o comportamento dinâmico da densidade de fluxo de calor através de uma parede de duas camadas (camada de tijolos e isolamento) é o objetivo do procedimento de identificação baseado em modelo neural deste trabalho. As entradas do sistema são: a temperatura interna ($u_1(t)$) e a temperatura externa ($u_2(t)$) da parede. O sinal de saída $y(t)$ consiste da densidade de fluxo de calor através de uma parede (Bloem, 1994). Neste caso, os dados (1680 amostras) usados no procedimento de identificação estão disponíveis na base de dados DaISy (De Moor, 2006). É importante mencionar que o sistema térmico, avaliado neste trabalho, é um sistema teste para métodos de identificação. Detalhes do sistema

térmico podem ser encontrados em Bloem (1994) e De Moor (2006).

3.2. Análise dos resultados de identificação

O procedimento de análise da identificação através de RN-RBF é dividido nas seguintes etapas: (i) escolha da estrutura da RN-RBF para representar o sistema a ser identificado, (ii) otimização da RN-RBF usando agrupamento de dados e pseudo-inversa; (iii) estimação dos parâmetros do modelo matemático (*fase de estimação*), e (iv) a validação do modelo matemático (*fase de validação*).

Na procura de um modelo “ideal”, devem ser observadas as seguintes questões: se o modelo encontrado é flexível o suficiente e se o modelo é desnecessariamente complexo, ou em outras palavras, se o modelo é confiável para o propósito desejado. Existem diversas maneiras de se realizar a validação de um modelo. Neste trabalho, o índice de desempenho utilizado é o erro médio quadrático (*Mean Squared Error*, MSE), ou seja, a meta é obter um menor MSE.

Para identificação foi escolhido um modelo matemático para a representação do processo. A estrutura de modelo testada é o modelo NARX (*Nonlinear AutoRegressive with eXogenous inputs*) em concepção série-paralela com 2 entradas ($[u_1(t-1); u_2(t-1)]$) e 1 saída estimada, $\hat{y}(t)$, para a densidade de fluxo de calor.

Na Tabela 1 é apresentado o resultado para o projeto do filtro de Kalman (executado apenas em uma época de aprendizado) para o treinamento da RN-RBF. Os limites para o espaço de busca utilizados na entropia cruzada foram $[LI, LS] = [0,1; 200]$ para as variáveis da diagonal principal de P , Q e R . Foram usadas as amostras de 1 a 840 na fase de estimação da RN-RBF, e as amostras de 841 a 1680 na fase de validação.

O melhor resultado obtido pela entropia cruzada para o projeto do filtro de Kalman usado na RN-RBF foi $P=169,7752$, $Q=102,8381$ e $R=95,1249$ usando uma RN-RBF com 3 funções Gaussianas (grupos) na camada oculta.

Tabela 1. Resultados das simulações para previsão densidade de fluxo de calor através de uma parede de duas camadas.

Gaussianas	Estimação	Validação
	MSE ($\times 10^{-3}$)	MSE ($\times 10^{-3}$)
2	0,8172	0,1250
3	0,4287	0,5134
4	0,4359	0,5135
5	0,4774	0,5134

Nas Figs. 1 e 2 são apresentados o sinal de erro e saída estimada para a densidade de fluxo de calor através de uma parede de duas camadas, respectivamente, para os melhores resultados apresentados na Tabela 1, estes obtidos pela RN-RBF com 3 funções Gaussianas na camada oculta. Nota-se que a RN-RBF apresentou um índice de desempenho

apropriado para modelar o comportamento dinâmico da densidade de fluxo de calor obtendo $MSE(\times 10^{-3}) = 0,4287$ na estimação e $MSE(\times 10^{-3}) = 0,5134$ na validação.

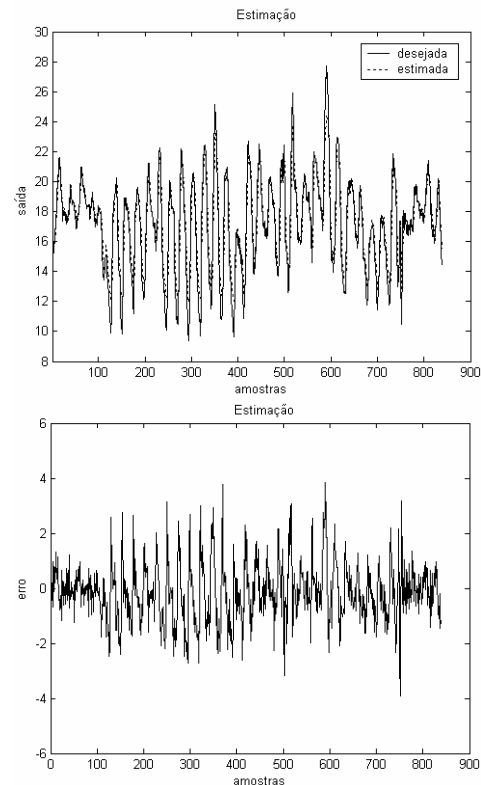


Fig. 1. Saída estimada durante a estimação para a densidade de fluxo de calor usando RN-RBF (3 funções Gaussianas).

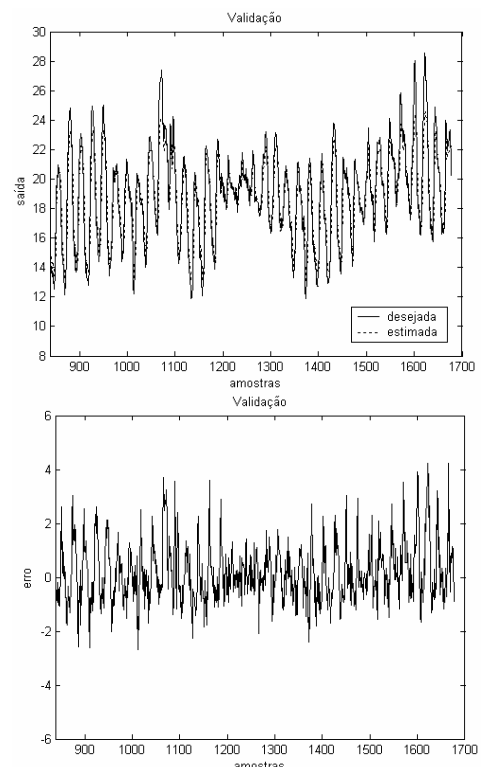


Fig. 2. Saída estimada durante a validação para a densidade de fluxo de calor usando RN-RBF (3 funções Gaussianas).

4 Conclusão

As RNs-RBF têm como uma das principais características agirem como se fossem aproximadores universais de funções, podendo ser utilizadas com êxito para identificar sistemas complexos. Este trabalho apresentou uma abordagem de RN-RBF baseada em treinamento por filtro de Kalman híbrido com entropia cruzada.

O filtro de Kalman possibilitou lidar com informações ruidosas do sistema dinâmico, utilizando informações matemáticas para estimar o estado atual do ambiente no qual está inserido. Uma dificuldade no projeto do filtro de Kalman foi o projeto das matrizes P , Q e R . Neste contexto, neste trabalho, utilizou-se a técnica de entropia cruzada para otimizar os termos da diagonal principal das matrizes de P , Q e R . Os resultados de empregar filtro de Kalman com entropia cruzada no projeto da RN-RBF foram satisfatórios em termos do critério de desempenho MSE.

Referências Bibliográficas

- Birgmeier, M. (1995). A fully Kalman-trained radial basis function network for nonlinear speech modeling, *Proceedings of IEEE International Conference on Neural Networks*, Vol. 1, pp. 259-264.
- Bloem, J. (ed.) (1994). System identification competition, benchmark tests for estimation methods of thermal characteristics of buildings and building components, Joint Research Centre, Ispra, Italy.
- Chen, S., Cowan, C. e Grant, P. (1991). Orthogonal least squares learning algorithm for radial basis function networks, *IEEE Transactions on Neural Networks*, **2**(2): 302-309.
- Côrrea, L. G. C. (2005). Inferência da qualidade de produtos de destilação utilizando redes neurais artificiais e filtro de Kalman estendido, *Dissertação de mestrado*, Programa de Pós-Graduação em Engenharia Elétrica, PUC-Rio, Rio de Janeiro, RJ.
- De Moor B. L. R. (ed.) (2006). *DaISy: Database for the identification of systems*, Department of Electrical Engineering, ESAT/SISTA, K.U.Leuven, Belgium, URL: <http://homes.esat.kuleuven.be/~smc/daisy/>, 10/06/2006.
- Haykin, S. (1996). *Neural networks*, 2nd edition, Prentice-Hall, Upper Saddle River, NJ, USA.
- Haykin, S. (2001). *Kalman filtering and neural networks*, John Wiley & Sons, Inc, NY, USA.
- Kalman, R. E. (1960). A new approach to linear filtering and prediction problems, *Transactions of the ASME – Journal of Basic Engineering*, **83**: 35-45.
- Kuchner, H. J. (1967). Approximations to optimal nonlinear filters, *IEEE Transactions on Automatic Control*, **12**(5): 546-556.
- Lieber, D. e Rubinstein, R. Y. (1999). Rare-event estimation via cross-entropy and importance sampling, *Technion*, Manuscript.
- Puskorius, G. V. e Feldkamp, L. A. (1994). Neurocontrol of nonlinear dynamical systems with Kalman filter trained recurrent networks, *IEEE Transactions on Neural Networks*, **5**(2): 279-297.
- Rubinstein, R. Y. (1997). Optimization of computer simulation models with rare events, *European Journal of Operations Research*, **99**: 89-112.
- Rubinstein, R. Y. (1999). The cross-entropy method for combinatorial and continuous optimization, *Methodology and Computing in Applied Probability*, **2**, 127-190.
- Simon, D. (2002). Training radial basis neural networks with the extended Kalman filter, *Neurocomputing*, **48**(1-4): 455-475.
- Xiong, S. -S. e Zhou, Z.-Y. (2003). Neural filtering of colored noise based on Kalman filter structure, *IEEE Transactions on Instrumentation and Measurement*, **52**(3): 742-747.
- Zhang, L. e Luh, P. B. (2005). Neural network-based market clearing price prediction and confidence interval estimation with an improved extended Kalman filter method, *IEEE Transactions on Power Systems*, **20**(1): 59-66.