NEURAL NETWORK IMPLEMENTATION OF ROBUST KALMAN PREDICTORS

Dušan KROKAVEC, Anna FILASOVÁ
Department of Cybernetics and Artificial Intelligence, Faculty of Electrical Engineering and Informatics,
Technical University of Košice, Letná 9, 042 00 Košice,
tel. +421 55 602 2564, E-mail: krokavec@tuke.sk, tel.+421 55 602 4128, E-mail: filasova@tuke.sk

SUMMARY

The purpose of the paper is to present an algorithm to solve the optimization tasks concerning with the robust discrete-time Kalman predictor design as well as the problem formulation for neural network implementation of robust Kalman predictor, where dual heuristic programming is used for realization. This application can be considered as a task concerned the class of problems referred to as reinforcement learning algorithms and is based on the existence of a complete model of the environment and the predicted error.

***Keywords**: **s**tate estimation, Kalman predictor, heuristic dynamic programming.*

1. INDRODUCTION

Searching for a control law for given model it is necessary to keep in mind that this model is far from perfect. This leads to the so-called robustness analysis of the plant and controllers. Robustness of the systems says nothing more than that the stability (or another global goal) of the system will stand against perturbation. The classic approach to this problem was LQG theory. In that approach the uncertainty is modeled as a white-noise Gaussian added as an extra input (vector) to the system. Because size of the error may be relative to the size of the inputs the parameter uncertainty can be better modeled as a disturbance in system, taking values in some range. The goal is to know the effect of the "worst" disturbance in the prescribed parameter range and if this disturbance cannot destabilize the system then we are certain that system is stabilized.

The discrete-time Kalman predictor design can be cast as an optimization problem that involve matrix equation, where this equation have the form of discrete or algebraic Riccati equation. In general, optimization problem solving need methods that can capture high order complexity and uncertainty of the system. One class of this method is Adaptive Critic Design (ACD). ACD approximate dynamic programming for optimal decision making in noisy, non-stationary or non-linear environments (heuristic dynamic programming). Using neural network a typical ACD include action, critic and model modules. Each module can be a neural network or, alternatively, any differentiable system. Heuristic dynamic programming is so a neural network approach to solve Bellman equation, where in Kalman prediction at least two neural nets are needed – one for functioning as the predictor gain (action), one used to train the predictor gain (critic) and a third could be trained to copy the error model. Good knowledge of the derivatives of an optimization criterion is a prerequisite to find a solution. Dual heuristic dynamic programming have an important advantage since its critic module produces a representation for parameter derivatives being explicitly trained on them.

The paper present new algorithms to solve the optimization tasks concerning with the robust discrete-time Kalman predictor design, where dual heuristic programming is used for realization The most applicable publications which have dealt with the above mentioned problem are presented in this paper References.

DISCRETE-TIME ROBUST KALMAN PREDICTOR

In general, a discrete-time stochastic uncertain multivariable system can be considered as

$$\mathbf{x}(i+1) = (\mathbf{F} + \Delta\mathbf{F}(i))\mathbf{x}(i) + \mathbf{G}\mathbf{u}(i) + \mathbf{v}(i) \tag{1}$$

$$\mathbf{y}(i) = (\mathbf{C} + \Delta\mathbf{C}(i))\mathbf{x}(i) + \mathbf{o}(i) \tag{2}$$

vectors $\mathbf{x}(i) \in \mathbf{R}^{n}$, $\mathbf{u}(i) \in \mathbf{R}^{r}$, $\mathbf{y}(i) \in \mathbf{R}^{m}$, system matrices $\mathbf{F} \in \mathbf{R}^{n\times n}$, $\mathbf{G} \in \mathbf{R}^{n\times r}$, $\mathbf{C} \in \mathbf{R}^{m\times n}$, are finite valued ones, and $\Delta\mathbf{F}(i) \in \mathbf{R}^{n\times n}$, $\Delta\mathbf{H}(t) \in \mathbf{R}^{n\times r}$ are unknown matrices which represents time-varying parametric uncertainties. It is assumed that considered uncertainty matrices to be of the form

$$\begin{bmatrix} \Delta\mathbf{F}(i) \\ \Delta\mathbf{C}(i) \end{bmatrix} = \begin{bmatrix} \mathbf{N}_1 \\ \mathbf{N}_2 \end{bmatrix} \mathbf{H}(i)\mathbf{M}\,, \tag{3}$$

$$\mathbf{H}^T(i)\mathbf{H}(i) \le \mathbf{I}, \tag{4}$$

where $\mathbf{N}_1 \in \mathbf{R}^{h\times n}$, $\mathbf{N}_2 \in \mathbf{R}^{h\times r}$, $\mathbf{M} \in \mathbf{R}^{n\times h}$ are known constant matrices which specify which elements of the nominal matrices $\mathbf{F}$ and $\mathbf{C}$ are affected by uncertain parameter matrix $\mathbf{H}$(i). The process and output noise random sequences are Gaussian with covariance $\mathbf{Q}_0 \in \mathbf{R}^{n\times n}$, $\mathbf{R}_0 \in \mathbf{R}^{m\times m}$, $\mathbf{S}_0 \in \mathbf{R}^{n\times m}$.

For such an system (1), (2), which must be stabilizable and detectable, given an initial estimate of the state $\mathbf{x}_e(0) = \mathbf{x}_0$ the robust Kalman predictor is defined as

$$\mathbf{x}_e(i+1) = \mathbf{F}\mathbf{x}_e(i) + \mathbf{G}\mathbf{u}(i) + \mathbf{K}(i)(\mathbf{y}(i) - \mathbf{y}_e(i)) \tag{5}$$

$$\mathbf{y}_e(i) = \mathbf{C}\mathbf{x}_e(i) \tag{6}$$

The actual error can be written as

$$\begin{aligned}\mathbf{e}(i+1) = \mathbf{x}(i+1) - \mathbf{x}_e(i+1) = \mathbf{v}(i) - \mathbf{K}(i)\mathbf{o}(i) + \\ +(\mathbf{F} - \mathbf{K}(i)\mathbf{C})\mathbf{e}(i) + (\Delta\mathbf{F}(i) - \mathbf{K}(i)\Delta\mathbf{C}(i))\mathbf{x}(i)\end{aligned} \tag{7}$$

The error expectation is zero and covariance of actual error is given by the equation

$$\begin{aligned}\mathbf{P}(i+1) = (\mathbf{F} - \mathbf{K}(i)\mathbf{C})\mathbf{P}(i)(\mathbf{F} - \mathbf{K}(i)\mathbf{C})^T + \mathbf{Q}_0 + \\ +\mathbf{K}(i)\mathbf{R}_0\mathbf{K}^T(i) - \mathbf{K}(i)\mathbf{S}_0^T - \mathbf{S}_0\mathbf{K}^T(i) + \\ +(\Delta\mathbf{F}(i) - \mathbf{K}(i)\Delta\mathbf{C}(i))\mathbf{W}^2(i)(\Delta\mathbf{F}(i) - \mathbf{K}(i)\Delta\mathbf{C}(i))^T\end{aligned} \tag{8}$$

where $\mathbf{W}^2(i) \in \mathbf{R}^{n\times n}$ is covariance of $\mathbf{x}(i) \in \mathbf{R}^n$.

It is straightforward to verify that yields

$$\begin{aligned}\mathbf{A}\mathbf{A}^T - \mathbf{A} - \mathbf{A}^T + \mathbf{I} = (\mathbf{I} - \mathbf{A})(\mathbf{I} - \mathbf{A})^T, \\ \mathbf{A}\mathbf{A}^T < (\mathbf{I})(\mathbf{I})^T - \mathbf{I} + \mathbf{A} + \mathbf{A}^T = \mathbf{A} + \mathbf{A}^T.\end{aligned} \tag{9}$$

Denoting the uncertainty elements of (8) as **L,** (9) implies that

$$\begin{aligned}\mathbf{L} = (\Delta\mathbf{F}(i) - \mathbf{K}(i)\Delta\mathbf{C}(i))\mathbf{W}^2(i)(\Delta\mathbf{F}(i) - \mathbf{K}(i)\Delta\mathbf{C}(i))^T \\ < (\Delta\mathbf{F}(i) - \mathbf{K}(i)\Delta\mathbf{C}(i))\mathbf{W}(i) + \\ +\mathbf{W}(i)(\Delta\mathbf{F}(i) - \mathbf{K}(i)\Delta\mathbf{C}(i))^T\end{aligned} \tag{10}$$

Since immediately follows from (3) and (10)

$$\begin{aligned}\mathbf{L} < \tfrac{\alpha}{\alpha}(\mathbf{N_1} - \mathbf{K}(i)\mathbf{N_2})\mathbf{H}(i)\mathbf{M}\mathbf{W}(i) + \\ +\mathbf{W}(i)\mathbf{M}^T\mathbf{H}^T(i)\tfrac{\alpha}{\alpha}(\mathbf{N_1} - \mathbf{K}(i)\mathbf{N_2})^T\end{aligned} \tag{11}$$

using condition (4) and the identity

$$\mathbf{A}\mathbf{B}^T + \mathbf{B}\mathbf{A}^T \le \mathbf{A}\mathbf{A}^T + \mathbf{B}\mathbf{B}^T \tag{12}$$

(11) may be rewritten as

$$\begin{aligned}\mathbf{L} < \varepsilon\mathbf{W}(i)\mathbf{M}^T\mathbf{H}^T(i)\mathbf{H}(i)\mathbf{M}\mathbf{W}(i) + \\ +\tfrac{1}{\varepsilon}(\mathbf{N_1} - \mathbf{K}(i)\mathbf{N_2})(\mathbf{N_1} - \mathbf{K}(i)\mathbf{N_2})^T \le \\ \le \varepsilon\mathbf{W}(i)\mathbf{M}^T\mathbf{M}\mathbf{W}(i) + \\ +\tfrac{1}{\varepsilon}(\mathbf{N_1} - \mathbf{K}(i)\mathbf{N_2})(\mathbf{N_1} - \mathbf{K}(i)\mathbf{N_2})^T\end{aligned} \tag{13}$$

or equivalently

$$\begin{aligned}\mathbf{L} < \mathbf{Q}^* + \mathbf{K}(i)\mathbf{R}^*\mathbf{K}^T(i) - \\ -\mathbf{K}(i)\mathbf{S}^{*T} - \mathbf{S}^*\mathbf{K}^T(i) + \mathbf{W}(i)\mathbf{M}^*\mathbf{W}(i)\end{aligned} \tag{14}$$

where $0 < \varepsilon = \alpha^2 \ll 1$ and

$$\begin{aligned}\mathbf{Q}^* = \tfrac{1}{\varepsilon}\mathbf{N}_1\mathbf{N}_1^T, \quad \mathbf{R}^* = \tfrac{1}{\varepsilon}\mathbf{N}_2\mathbf{N}_2^T, \\ \mathbf{S}^* = \tfrac{1}{\varepsilon}\mathbf{N}_1\mathbf{N}_2^T, \quad \mathbf{M}^* = \varepsilon\mathbf{M}^\mathbf{T}\mathbf{M}.\end{aligned} \tag{15}$$

The upper bound of actual error covariance (8) is

$$\begin{aligned}\mathbf{P}(i+1) = (\mathbf{F} - \mathbf{K}(i)\mathbf{C})\mathbf{P}(i)(\mathbf{F} - \mathbf{K}(i)\mathbf{C})^T + \\ +\mathbf{Q}(i) + \mathbf{K}(i)\mathbf{R}\mathbf{K}^T(i) - \mathbf{K}(i)\mathbf{S}^T - \mathbf{S}\mathbf{K}^T(i)\end{aligned} \tag{16}$$

where

$$\begin{aligned}\mathbf{Q}(i) = \mathbf{Q}_0 + \mathbf{Q}^* + \mathbf{W}(i)\mathbf{M}^*\mathbf{W}(i), \\ \mathbf{R} = \mathbf{R}_0 + \mathbf{R}^*, \quad \mathbf{S} = \mathbf{S}_0 + \mathbf{S}^*.\end{aligned} \tag{17}$$

Value of (16) can be minimized if

$$\mathbf{K}(i) = (\mathbf{F}\mathbf{P}(i)\mathbf{C}^T + \mathbf{S})(\mathbf{R} + \mathbf{C}\mathbf{P}(i)\mathbf{C}^T)^{-1} \tag{18}$$

and the error covariance matrix is a solution of the Riccati equation

$$\mathbf{P}(i+1) = \mathbf{Q} + \mathbf{F}\mathbf{P}(i)\mathbf{F}^T - \mathbf{K}(i)(\mathbf{C}\mathbf{P}(i)\mathbf{F}^T + \mathbf{S}^T) \tag{19}$$

The design progresses forward in time from time point *i*=0 and is looking for the steady-state solution.

KALMAN PREDICTOR ERROR MODEL

It is presumed, that predictor error is a solution of the equation (7) and the follow error model

$$\begin{aligned}E\{\mathbf{e}^T(i+1)\mathbf{e}(i)\} = \\ E\{\mathbf{e}^T(i)(\mathbf{F} - \mathbf{K}(i)\mathbf{C})^T\mathbf{e}(i) + (\mathbf{v}(i) - \mathbf{K}(i)\mathbf{o}(i))^T\mathbf{e}(i) + \\ +\mathbf{x}^T(i)(\Delta\mathbf{F}(i) - \mathbf{K}(i)\Delta\mathbf{C}(i))^T\mathbf{e}(i)\} = \\ \mathbf{e}^T(i)(\mathbf{F} - \mathbf{K}(i)\mathbf{C})^T\mathbf{e}(i) = E\{\mathbf{e}^T(i)\mathbf{e}(i+1)\}\end{aligned} \tag{20}$$

can be obtained. Note that the deterministic part of this model imply the error state-space equations

$$\mathbf{e}(i+1) = \mathbf{F}^T\mathbf{e}(i) + \mathbf{C}^T\mathbf{q}(i) \tag{21}$$

$$\mathbf{q}(i) = -\mathbf{K}(i)\mathbf{e}(i) \tag{22}$$

which can be considered as a linear quadratic optimal controlled dual system with performance index

$$\begin{aligned}J = \sum_{i=0}^{\infty}(\mathbf{e}^T(i)\mathbf{Q}(i)\mathbf{e}(i) + \mathbf{e}^T(i)\mathbf{S}\mathbf{q}(i) + \\ +\mathbf{q}(i)\mathbf{S}^T\mathbf{e}(i) + \mathbf{q}^T(i)\mathbf{R}\mathbf{q}(i)) = \sum_{i=1}^{\infty} r(\mathbf{e}(i), \mathbf{q}(i))\end{aligned} \tag{23}$$

DUAL HEURISTIC PROGRAMMING

The system studied is on the form of discrete-time state-space description, the performance index is (23), $\mathbf{e}(i+1) = f(\mathbf{e}(i),\mathbf{q}(i))$ is the state equation (21), $\mathbf{q}(i) = g(\mathbf{e}(i))$ (22) is controlled inputs and $V(\mathbf{e}(i))$ is the Lyapunov function, i.e.

$$f(\mathbf{e}(i),\mathbf{q}(i)) = \mathbf{F}^T\mathbf{e}(i) + \mathbf{C}^T\mathbf{q}(i) = \mathbf{e}(i+1) \tag{24}$$

$$\begin{gathered} r(\mathbf{e}(i),\mathbf{q}(i)) = \mathbf{e}^T(i)\mathbf{Q}(i)\mathbf{e}(i) + \\ +\mathbf{q}^T(i)\mathbf{R}\mathbf{q}(i) + \mathbf{e}^T(i)\mathbf{S}\mathbf{q}(i) + \mathbf{q}^T(i)\mathbf{S}^T\mathbf{e}(i) \end{gathered} \tag{25}$$

$$g(\mathbf{e}(i)) = -\mathbf{K}(i)\mathbf{e}(i) = \mathbf{q}(i) \tag{26}$$

$$V(\mathbf{e}(i)) = \mathbf{e}^T(i)\mathbf{P}(i)\mathbf{e}(i) \tag{27}$$

One way to search for the optimal parameters is to employ a gradient algorithm.

The Pontryagin minimum principle implies, if there is no bounds on $\mathbf{q}$(i), the minimizing $\mathbf{q}$(i) must be such, that

$$\begin{gathered} \frac{\partial V(\mathbf{e}(i))}{\partial g(\mathbf{e}(i))} = \\ = \frac{\partial r(\mathbf{e}(i),\mathbf{q}(i))}{\partial g(\mathbf{e}(i))} + \frac{\partial V(\mathbf{e}(i+1))}{\partial \mathbf{e}(i+1)} \frac{\partial f(\mathbf{e}(i),\mathbf{q}(i))}{\partial g(\mathbf{e}(i))} = 0 \end{gathered} \tag{28}$$

In this sense the target for an action network minimization can be defined as zero and the network output error is

$$\begin{gathered} e_a(i) = 0 - \frac{\partial V(\mathbf{e}(i))}{\partial g(\mathbf{e}(i))} = \\ -\left(\frac{\partial r(\mathbf{e}(i),\mathbf{q}(i))}{\partial g(\mathbf{e}(i))} + \frac{\partial V(\mathbf{e}(i+1))}{\partial \mathbf{e}(i+1)} \frac{\partial f(\mathbf{e}(i),\mathbf{q}(i))}{\partial g(\mathbf{e}(i))} \right) \end{gathered} \tag{29}$$

Using the criterion

$$W_a(i) = \tfrac{1}{2}\mathbf{e}_a^T(i)\mathbf{e}_a(i) = \tfrac{1}{2}\sum_{h=1}^{M} e_{ah}^2(i) \tag{30}$$

for the action neural network training, a steepest-descent discrete gradient method, based on error back-propagation algorithm, can be applied to solve this minimization problem, i.e.

$$\begin{gathered} \Delta w_{rs}(i) = -\mu_a \frac{\partial W_a(i)}{\partial w_{rs}(i)} = \\ = -\mu_a \left(\frac{\partial r(i)}{\partial \mathbf{q}(i)} + \frac{\partial V(i+1)}{\partial \mathbf{e}(i+1)} \frac{\partial \mathbf{e}(i+1)}{\partial \mathbf{q}(i)} \right) \frac{\partial \mathbf{q}(i)}{\partial w_{rs}(i)} = \\ = -\mu_a \sum_{k=1}^{M} \left(\frac{\partial r(i)}{\partial u_k(i)} + \sum_{j=1}^{N} \frac{\partial V(i+1)}{\partial e_j(i+1)} \frac{\partial e_j(i+1)}{\partial u_k(i)} \right) \frac{\partial q_k(i)}{\partial w_{rs}(i)} \end{gathered} \tag{31}$$

where $\frac{\partial e_j(i+1)}{\partial u_k(i)}$, $\frac{\partial W(i+1)}{\partial e_j(i+1)}$ and $\frac{\partial r(i)}{\partial u_k(i)}$ are calculated from analytical equation of the error model, approximated by the critic network, and calculated as a derivative of performance index, respectively. Variable M and N designate the number of state and input variables.

CRITIC NEURAL NETWORK TRAINING

Heuristic dynamic programming has a critic network that estimates values of the function V($\mathbf{e}$(i)). The critic neural network is trained using the assumption of the optimal response

$$V(\mathbf{e}(i)) \doteq V(\mathbf{e}(i+1)) + r(\mathbf{e}(i),\mathbf{q}(i)) \tag{32}$$

The critic is trained forward in time, which is of the great importance for real-time operation.

When heuristic programming is used, the critic network tries to minimize the following error measure over time

$$W_c = \sum_{i=0}^{\infty} W_c(i) = \sum_{j=0}^{\infty} (c(i) - c^0(i))^2 \tag{33}$$

where desired output of the critic neural network in the time point *i* is the value of the function (31), i.e.

$$c^0(i) = V(\mathbf{e}(i+1)) + r(\mathbf{e}(i),\mathbf{q}(i)) \tag{34}$$

There is used value of $V(\mathbf{e}(i+1))$ at the time-point i+1 and as the output of critic neural network is taken into account the value $V(\mathbf{e}(i))$ from the last-but-one step of iteration, i.e.

$$c(i) = V(\mathbf{e}(i)) \tag{35}$$

and the optimization procedure is given by

$$\Delta w_{rs}(i) = -\mu_c W_c(i) \frac{\partial V(\mathbf{e}(i))}{\partial w_{rs}(i)} \tag{36}$$

where μ is a positive learning rate.

Dual heuristic dynamic programming and its action dependent form have a critic network that estimates the derivatives of cost function with respect to the prediction error vector $\mathbf{e}$(i), which gives for the *j*-th desired output of the critic neural network

$$c_j^o(i) = \frac{\partial V(i)}{\partial e_j(i)} = \frac{\partial r(\mathbf{e}(i),\mathbf{q}(i))}{\partial e_j(i)} +$$

$$+\frac{\partial r(\mathbf{e}(i),\mathbf{q}(i))}{\partial \mathbf{q}(i)}\frac{\partial \mathbf{q}(i)}{\partial e_j(i)} + \frac{\partial V(\mathbf{e}(i+1))}{\partial \mathbf{e}(i+1)}\frac{\partial \mathbf{e}(i+1)}{\partial e_j(i)} =$$

$$= \frac{\partial r(i)}{\partial e_j(i)} + \sum_{k=1}^{M}\frac{\partial r(i)}{\partial q_k(i)}\frac{\partial q_k(i)}{\partial e_j(i)} + \qquad (37)$$

$$+\sum_{h=1}^{N}\frac{\partial V(i+1)}{\partial e_h(i+1)}\frac{\partial e_h(i+1)}{\partial e_j(i)} +$$

$$\sum_{k=1}^{M}\sum_{h=1}^{N}\frac{\partial V(i+1)}{\partial e_h(i+1)}\frac{\partial e_h(i+1)}{\partial q_k(i)}\frac{\partial q_k(i)}{\partial e_j(i)}$$

where $\frac{\partial e_j(i+1)}{\partial u_k(i)}$, $\frac{\partial W(i+1)}{\partial e_j(i+1)}$, $\frac{\partial r(i)}{\partial u_k(i)}$, $\frac{\partial r(i)}{\partial e_j(i)}$ are calculated from analytical equation of the error model, approximated by the critic network, and calculated as a derivative of performance index, respectively. The value $\frac{\partial u_k(i)}{\partial e_j(i)}$ is given as the product of synaptic weights on the path from the *j*-th input to *k*-th output of the action neural network.

The training criterion for critic neural network can be defined as

$$W_c(i) = \tfrac{1}{2}\sum_{j=1}^{N}(c_j(i) - c_j^0(i))^2 \qquad (38)$$

and the neural network optimization procedure is given by

$$\Delta w_{rs}(i) = -\mu_c \frac{\partial W_c(i)}{\partial w_{rs}(i)} \qquad (39)$$

It is evident that the basic strategy to update the networks can be given by the straight application of (30), (38). The better critic neural network approximate criterion the better the action neural network will approximate an optimal control.

In dual heuristic programming a target (desired output) is needed for training the critic network and this is typically calculated by running the critic network one more computational cycle to provide its next-in-time output, and then use this value to compute the target for the present-time cycle. The error term is calculated and the critic network update is performed in the usual way. Since the critic network that calculates the target is changing with each update, it provides a moving target for critic neural network training.

CONCLUDING REMARKS

The paper present some background material on the robust discrete time Kalman predictor, an overview of the dual heuristic programming problem and a survey of techniques considered from the point of feed-forward multilayer perceptron and neural network training.

Applications can be considered as a task concerned the class of problems referred to as reinforcement learning algorithms. Reinforcement learning is a general way to formulate complex learning problems. The goal of the system is to maximize a long terms sum of an instantaneous reward (provided by the teacher). It is a decision process based on system environment simulation and in its extremum form it only requires that the teacher can provide a measure of success.

Presented new application of dual heuristic programming principle for discrete-time Kalman state estimation, based on the existence of a complete model of the environment and the dual-system model of prediction error, involved back-propagation utilities with system response parameterization. This approximation of the gradient algorithms for parameter updating in the sense of the mean value for given training set is a basic one for implementation of presented tasks using adaptive critic design for neuro-control, which is suitable for learning in noisy and non-stationary environments.

ACKNOWLEDGEMENTS

The work presented in this paper was supported by Grant Agency of Ministry of Education and Academy of Science of Slovak Republic VEGA under Grant No. 1/9028/02. The paper was presented at the 5th International Conference Digital Signal Processing and Multimedia Communication DSP-MCOM 2001, November 27-29, 2001, Košice, Slovakia.

REFERENCES

[1] BELLMAN, R. E. - KALABA, R.: *Dynamic Programming and Modern Control Theory.* New York : Academic Press, 1965.

[2] A. CHICHOCKY, A. - UNBEHAUEN, R.: *Neural Networks for Optimization and Signal Processing*, New York: Wiley, 1993, ISBN 0-471-93010-5.

[3] FILASOVÁ, A. - KROKAVEC, D.: Duality of Kalman estimators and LQR controllers from the point of dual dynamic programming. In *Process control '98: proceeding of the 3rd scientific - technical conference with international participation ŘIP '98 June 3-6, 1998, Vol. 1* / S. Krejčí, B. Jakeš, J. Macháček, I. Taufer

(eds.). Kouty nad Desnou, Czech Republic, pp. 107-112. ISBN 80-7194-051-8
[4] FILASOVÁ, A. - KROKAVEC, D.: Heuristic dynamic programming for LQR controllers and Kalman estimators. In *Artificial intelligence in industry: proceedings of the scientific conference AIII '98 April 22-24, 1998.* / J. Sarnovský, P. Sinčák, M. Mach, M. Hatala (eds.). Stará Lesná, Slovak Republic, pp. 110-119. ISBN 80-88786-83-5.
[5] KROKAVEC, D. - FILASOVÁ, A.: *Optimal stochastic systems.* Košice, Elfa, 1999, ISBN 80-88964-03-2. (in Slovak)
[6] KROKAVEC, D. - FILASOVÁ, A.: Aplication of heuristic dynamic programming to dynamic system stabilization. In *State of art in computational intelligence*. / P. Sinčák, J. Vačšák, V. Kvasnička, R. Mesiar (eds.). Heidelberg : Physica Verlag, 2000, pp. 196-201. (Advances in Soft Computing). ISBN 3-7908-1322-2, ISSN 1615-3871.
[7] KROKAVEC, D. - FILASOVÁ, A.: Optimization in LQ control and Kalman prediction using neural network. In *Control system design: a preprints volume from the IFAC conference June 18-20, 2000.* / Š. Kozák, M. Huba (eds.). Bratislava, Slovak Republic, pp. 554-559.
[8] LANDELIUS, T. - KNUTSSON, H.: *Greedy adaptive critics for LQR problems.* Technical Report LiTH -ISY-R-1896. Linköping University, Sweden, 1996.
[9] PROKHOROV, D. V. - SANTIAGO, R. A. - WUNSCH, D. C.: Adaptive critic design: a case study for neurocontrol. *Neural Networks*, 1995, Vol. 8, No. 9, pp. 1367-1372. ISSN 0893-6080.
[10] PROKHOROV, D. - WUNSCH, D.: Adaptive critic design. *IEEE Transactions on Neural Networks*, 1997, Vol. 8, No. 5, pp. 997-1007. ISSN 1045-9227.
[11] WERBOS, P. J.: Consistency of HDP applied to a simple reinforcement learning problem. *Neural Networks*, 1990, Vol. 3, No. 2, pp. 179-189. ISSN 0893-6080.

BIOGRAPHY

Dušan Krokavec (prof., Ing. CSc.) was born in Ratková, Slovakia in 1945. He graduated in automation engineering from the Slovak Technical University in Bratislava in 1967 and received the PhD in technical cybernetics from the Slovak Technical University in Bratislava in 1982. From 1968 to 1971 he was a Research Assistant at the Research Institute of Automation and Mechanization in Nové Mesto n/Váhom, Slovakia. Since 1971 he has been with the Department of Cybernetics and Artificial Intelligence, Faculty of Electrical Engineering and Informatics, Technical University in Košice, promoted to an Associated Professor since 1984. In 1999 he was appointed Full Professor in automation and control. In the research field he is concerned with stochastic processes in dynamic systems, digital control systems and digital signal processing, robust control, dynamic system fault diagnosis, as well as with speech processing system application.

Anna Filasová (doc., Ing., CSc.) was born in Košice. She graduated in technical cybernetics from the Technical University in Košice in 1975 and received the PhD in technical cybernetics from the Technical University in Košice in 1993. From 1975 to 1999 she has been an Assistant Professor at the Department of Cybernetics and Artificial Intelligence, Faculty of Electrical Engineering and Informatics, Technical University in Košice. In 1999 she was appointed Associate Professor in automation and control from the Technical University in Košice. Her main research interests are in the areas of decentralized control, large-scale system optimization, and robust and predictive control.