



INŽENÝRSKÁ MECHANIKA 2005

NÁRODNÍ KONFERENCE

s mezinárodní účastí

Svratka, Česká republika, 9. - 12. května 2005

THE EFFECT OF SELECTED VARIABLES ON Q-LEARNING DURING THE CONTROL OF ASYNCHRONOUS ELECTRIC MOTOR

Tomáš Marada¹, Tomáš Březina²

Summary: *Asynchronous electric motor control task can be successfully solved using reinforcement learning based method called Q-learning. The main problem to solve is the convergence speed. Two-phase Q-learning can be used to speed up the learning process. Efficient prelearning phase uses mathematical model and next phase using for tutoring real system. This method can increase learning speed significantly. The paper is focused on presentation of random error effect of system variable observation, delay of interference action and response on jump torque onto convergence speed.*

1. Úvod

V poslední době je značná pozornost věnována metodám opakovaně posilovaného učení (reinforcement learning, RL). Vzhledem k tomu, že je poměrně obtížné získat množinu vzorů požadovaného chování regulátoru, což je nezbytná podmínka při učení s učitelem, nabízí se použití metod opakovaně posilovaného učení, které tuto množinu vzorů nevyžadují. Tyto metody mohou být také použity ke zlepšení kvality řízení asynchronního elektromotoru, případně ke snížení nároků na řídicí elektroniku.

Jedním z nejpopulárnějších a nejefektivnějších algoritmů opakovaně posilovaného učení bez modelu je Q-učení. Tohoto algoritmu je použito pro řízení asynchronního elektromotoru. Proces učení konvenčních architektur Q-učení je pro reálné úlohy řízení příliš pomalý. Použitelnou metodou na zrychlení učení, je rozdělení učení do dvou fází. Na fázi před-učení a fázi doučování. Fáze předučení vyžaduje výpočtový model, ale je velmi efektivní. Fáze doučování používá konvenčního Q-učení a předpokládá interakci s řízenou soustavou.

V příspěvku je studován vliv odolnosti strategie vůči náhodným chybám pozorování veličin soustavy, odolnost strategie vůči zpoždění akčního zásahu a odezva na skokový moment s lineárními a nelineárními mřížkami Q-funkce ve fázi před-učení.

¹ Tomáš Marada, Ing., Ph.D., VUT v Brně, FSI ÚAI, Technická 2, 616 69 Brno, marada@fme.vutbr.cz

² Tomáš Březina, Doc., Ing., RNDr., CSc., VUT v Brně, FSI ÚAI, Technická 2, 616 69 Brno, brezina@fme.vutbr.cz

2. Q-učení

Popis Q učení byl již publikován v mnoha člancích, například [1], a proto se jím již nebudeme zabývat. Implementace Q-funkce byla v našem případě provedena tabulkou. Stav prostředí s je chápán jako aritmetický vektor jistých veličin prostředí $\mathbf{s} = (s^1, \dots, s^n)$, a podobně akce u jako aritmetický vektor řídicích veličin $\mathbf{u} = (u^1, \dots, u^m)$. Necht' $\{g_j^i\}_{j=0}^{p_i}$, je posloupnost uzlových bodů rastru i -té veličiny prostředí z s^i taková, že $g_j^i < g_{j'}^i \Leftrightarrow j < j'$, $j, j' = 0, \dots, p_i$. Hodnotou indexu, který reprezentuje hodnotu i -té veličiny prostředí vzhledem k rastru, nazveme takový index j , pro který $g_j^i \leq s^i < g_{j+1}^i$. Zcela analogicky je zavedena reprezentace akce skládající se z řídicích veličin s uzlovými body rastru $\{h_j^i\}_{j=0}^{r_i}$. Funkce $Q(\mathbf{s}, \mathbf{u})$ je potom reprezentována tabulkou jako

$$Q(\mathbf{s}, \mathbf{u}) = q_{j^1 j^2 \dots j^{n+m}}, \quad (1)$$

kde j^i reprezentuje hodnotu i -té veličiny prostředí (resp. $i - n$ -té řídicí veličiny).

Přiměřenost aproximace funkce $Q(\mathbf{s}, \mathbf{u})$ tabulkou bude ovlivněna konkrétní volbou uzlových bodů rastru. Bude-li polí rastru málo, půjde zřejmě o příliš hrubou aproximaci, bude-li rastr příliš jemný, stane se neúnosným požadavek modifikací Q-hodnot nad značně vysokým počtem polí. V prvním případě je vyloučeno dosažení funkce $Q^*(\mathbf{s}, \mathbf{u})$ s dostatečnou přesností, v druhém případě toto není dosažitelné bez zdlouhavého výpočtu.

Dalším faktorem je to, že „kvalita“ strategie řízení bude zřejmě záviset i na velikosti a poloze (tj. uzlových bodech) jednotlivých polí rastru. Pro množinu stavů s požadavkem „jemnějších“ řídicích zásahů bude třeba „jemnějšího“ rastru, než pro množinu bez této podmínky.

Běžnou strategií při použití Q-učení je využívat k učení všech stavů prostředí, které postupně nastanou v průběhu jednotlivých pokusů. Tato strategie se nazývá učení pokusem a byla použita i v našem případě.

3. Aplikace

Jako modelu prostředí je použito známého modelu asynchronního elektromotoru v transformovaných souřadnicích $d, q, 0$. Tento model používá rovnice pro statorová napětí ve tvaru:

$$\begin{bmatrix} u_{qs} \\ u_{ds} \\ u_{0s} \end{bmatrix} = R_s \begin{bmatrix} i_{qs} \\ i_{ds} \\ i_{0s} \end{bmatrix} + \frac{d}{dt} \begin{bmatrix} \Psi_{qs} \\ \Psi_{ds} \\ \Psi_{0s} \end{bmatrix}, \quad (2)$$

rovnice pro rotorová napětí ve tvaru:

$$\begin{bmatrix} u_{qr} \\ u_{dr} \\ u_{0r} \end{bmatrix} = R_r \begin{bmatrix} i_{qr} \\ i_{dr} \\ i_{0r} \end{bmatrix} - \omega_r \begin{bmatrix} \Psi_{dr} \\ -\Psi_{qr} \\ 0 \end{bmatrix} + \frac{d}{dt} \begin{bmatrix} \Psi_{qr} \\ \Psi_{dr} \\ \Psi_{0r} \end{bmatrix}, \quad (3)$$

rovnice pro spřažené magnetické toky rotoru a statoru ve tvaru:

$$\begin{bmatrix} \Psi_{qs} \\ \Psi_{ds} \\ \Psi_{0s} \\ \Psi_{qr}^* \\ \Psi_{dr}^* \\ \Psi_{0r}^* \end{bmatrix} = \begin{bmatrix} L_{ls} + L_m & 0 & 0 & L_m & 0 & 0 \\ 0 & L_{ls} + L_m & 0 & 0 & L_m & 0 \\ 0 & 0 & L_{ls} & 0 & 0 & 0 \\ L_m & 0 & 0 & L_{lr}^* + L_m & 0 & 0 \\ 0 & L_m & 0 & 0 & L_{lr}^* + L_m & 0 \\ 0 & 0 & 0 & 0 & 0 & L_{lr}^* \end{bmatrix} \begin{bmatrix} i_{qs} \\ i_{ds} \\ i_{0s} \\ i_{qr}^* \\ i_{dr}^* \\ i_{0r}^* \end{bmatrix} \quad (4)$$

a rovnici pro elektromagnetický moment stroje ve tvaru:

$$M_e = \frac{3}{2} \frac{P}{2} \left[\left(\Psi_{qr}^* i_{dr}^* - \Psi_{dr}^* i_{qr}^* \right) \right] = \frac{3}{2} \frac{P}{2} (\Psi_{ds} i_{qs} - \Psi_{qs} i_{ds}). \quad (5)$$

Implementace modelu prostředí je provedena v prostředí Borland Delphi a k řešení používá lichoběžníkovou metodu integrace. Z modelu vystupuje aktuální regulační odchylka, ze které je vypočtena aktuální rychlost a zrychlení regulační odchylky.

4. Provedené experimenty

Experimenty se týkají vlivu odolnosti strategie vůči náhodným chybám pozorování veličin soustavy, odolnost strategie vůči zpoždění akčního zásahu a odezva na skokový moment s lineárními a nelineárními mřížkami Q-funkce ve fázi před-učení.

Při všech experimentech byly akce agenta brány z množiny $\{0Hz, 50Hz\}$, kde hodnoty „0Hz“ a „50Hz“ udávají požadovanou frekvenci na výstupu frekvenčního měniče. Této hodnoty je postupně dosaženo pomocí rampy frekvenčního měniče.

Posilovací funkce r byla definována jako:

$$r = \begin{cases} 0, & \text{pro } |e_k| < |e_{k-1}|, \text{ nebo } |e_k| \leq e_{\text{pož}} \\ -1, & \text{jinak} \end{cases} \quad (6)$$

kde e_k je aktuální regulační odchylka, e_{k-1} je předchozí regulační odchylka a $e_{\text{pož}}$ je požadovaná regulační odchylka.

Pro experimenty bylo použito hodnot parametrů modelu uvedených v tab. 1. Tyto byly získány identifikačními technikami z asynchronního elektromotoru.

$R_r [\Omega]$	$R_s [\Omega]$	$L_r [H]$	$L_s [H]$	$L_{rs} [H]$	$J [kg \cdot m^2]$
2.95	4.37	0.476	0.471	0.459	0.0015

Tab. 1.: Parametry asynchronního elektromotoru 4AP 90S-2

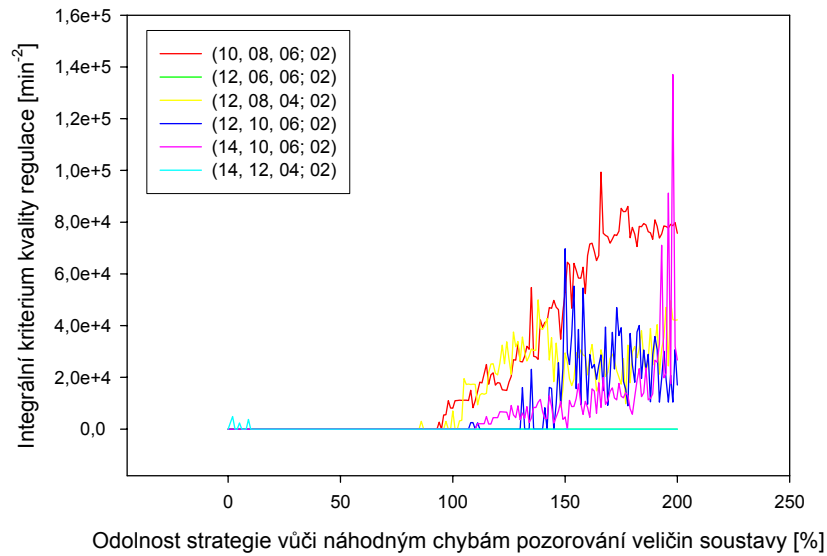
V provedených experimentech byla zjišťována kvalita regulace během 5000 kroků učení. Během každého kroku učení byl 50x proveden rozběh asynchronního elektromotoru z nulových otáček na hodnotu požadovaných otáček se zátěžným momentem $M_z = 2.5 Nm$. Požadované otáčky se měnily v rozsahu 250 min^{-1} až 2750 min^{-1} , aby byla vyzkoušena co nejširší oblast regulace. Kvalita regulace byla měřena integrálním kritériem kvality regulace.

V legendách následujících grafů udává první číslice počet dílků rastru regulační odchylky e , druhé číslo udává počet dílků rastru její rychlosti \dot{e} a třetí číslo udává počet dílků rastru jejího zrychlení \ddot{e} . Poslední číslo udává počet použitých akcí.

5. Fáze předučení s lineárními 3-D mřížkami Q-funkce

5.1. Odolnost strategie vůči náhodným chybám pozorování veličin soustavy

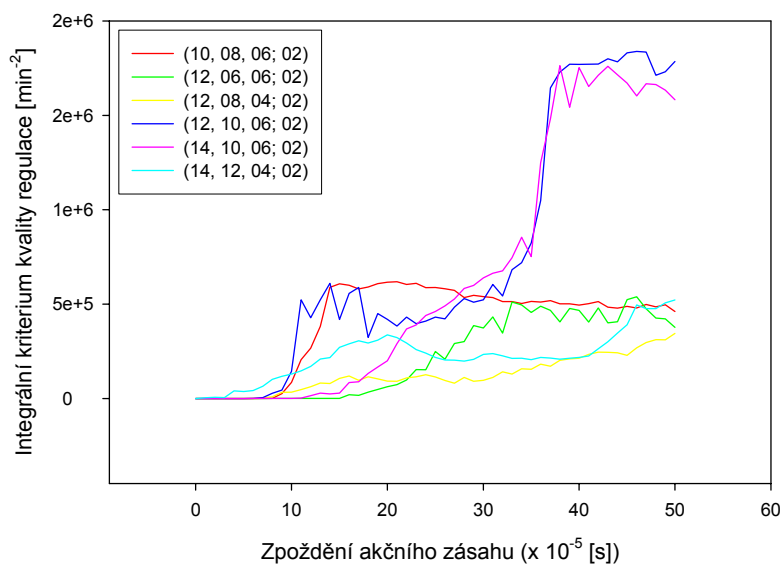
Výsledky testů jsou shrnuty na obr. 1. Chyby pozorování veličin soustavy byly zavedeny do všech veličin stavu soustavy. U mřížek byla zachována průměrná hodnota integrálního kritéria kvality regulace, až přibližně do 80 % úrovně chyb. Takto velká odolnost je s největší pravděpodobností zapříčiněna způsobem generování velikosti náhodných chyb. Tato procentuální velikost je totiž odvozena z aktuální hodnoty veličiny popisující stav prostředí (regulační odchylka, rychlost regulační odchylky, zrychlení regulační odchylky) a protože při řízení kolem ustálené polohy je tato velikost minimální můžeme si dovolit, až téměř 80 procentní chybu pozorování veličin soustavy. S dalším zvyšováním úrovně chyb pozorování veličin soustavy se integrální kritérium kvality regulace zvyšuje. Nejvyšší odolnost vůči náhodné chybě pozorování veličin soustavy vykazuje mřížka (14, 12, 4; 2).



Obr. 1.: Odolnost strategie vůči náhodným chybám pozorování veličin soustavy

5.2. Odolnost strategie vůči zpoždění akčního zásahu

Výsledky experimentů jsou shrnuty na obr. 2. U všech mřížek mají závislosti velmi podobný charakter. Nad hodnotou přibližně 10×10^{-5} [s] začíná prudce stoupat hodnota integrálního kritéria kvality regulace. Nejvíce vyniká mřížka (12, 6, 6; 2) u které se zlomová hodnota zpoždění akčního zásahu posouvá přibližně na hodnotu 16×10^{-5} [s] a která průběh stoupání hodnoty integrálního kritéria kvality regulace nemá tak strmý.

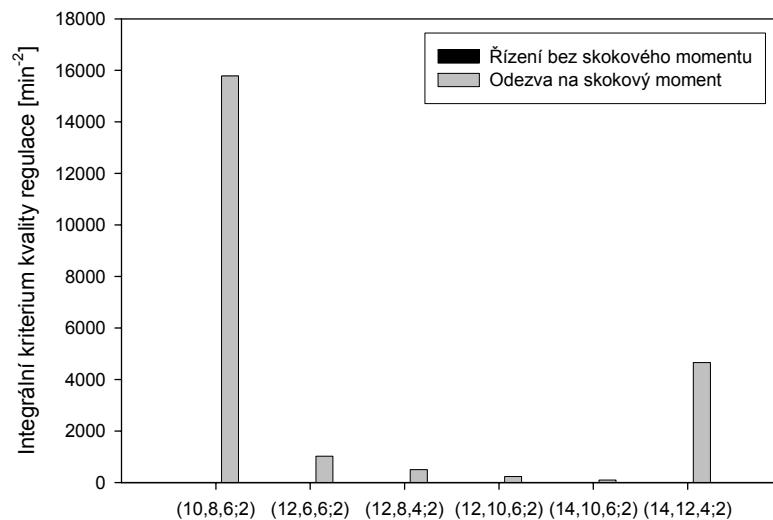


Obr. 2.: Odolnost strategie vůči zpoždění akčního zásahu

5.3. Odezva na skokový moment

Výsledky testů odezvy řízení na skokový moment jsou uvedeny na obr. 3. Experimenty probíhali následovně. Nejdříve byl proveden rozběh AM bez zátěžného momentu o délce trvání 9s. Poté byl motor zatížen zátěžným momentem $M_z = 5Nm$ a bylo měřeno integrální kritérium kvality regulace. V grafu první hodnota udává velikost integrálního kritéria kvality regulace bez skokového momentu a druhá hodnota udává velikost integrálního kritéria se skokovým momentem.

U všech testovaných mřížek byla hodnota integrálního kritéria kvality regulace bez skokového momentu rovna 0. Po zatížení skokovým momentem se nejlépe chovala mřížka (14, 10, 6; 2). Nejhuře se chovala mřížka (10, 8, 6; 2).



Obr. 3.: Odezva na skokový moment

5.4. Výběr vhodné lineární mřížky

Z posuzovaných lineárních mřížek 3-D stavu byla jako nejlepší zvolena mřížka (12, 8, 4; 2), protože vykazovala dobrou odolnost vůči náhodné chybě pozorování veličin soustavy, dobrou odolnost vůči zpoždění akčního zásahu a dobrou odezvu na skokový moment.

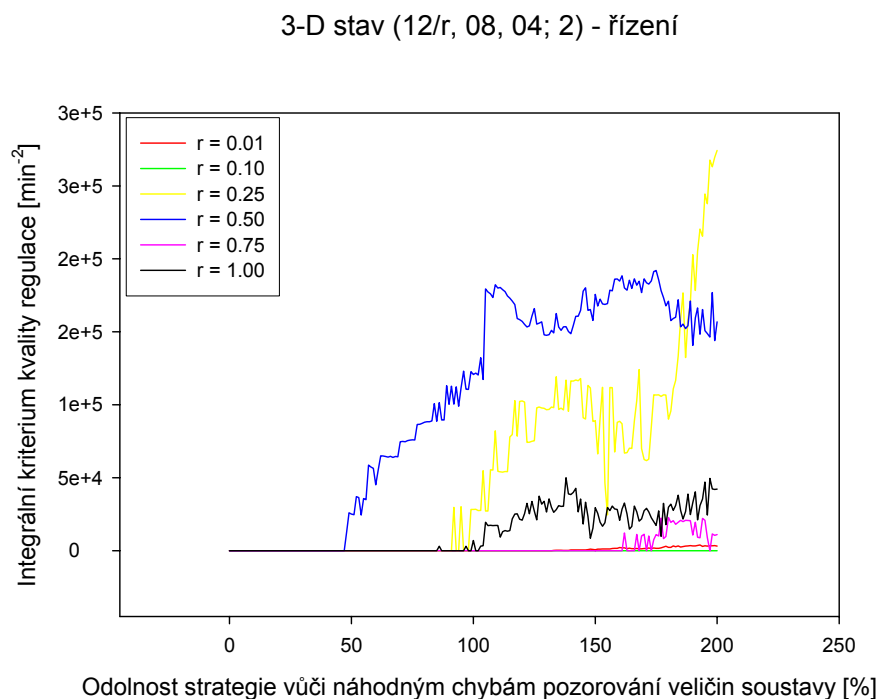
6. Fáze předučení s nelineárními 3-D mřížkami Q-funkce

Simulace byly navrženy s cílem ověřit, zda zmenšením vzdálenosti uzlových bodů nejbližších uzlovému bodu, který reprezentuje nulovou hodnotu veličiny, dojde ke zvýšení rozlišovací schopnosti strategie získané předučením v oblasti malých hodnot veličin a tím i ke zlepšení integrálního kritéria kvality regulace. K simulacím bylo použito strategie s 3-D

mřížkou (12, 8, 4; 2) která byla získána předchozími experimenty. Faktor nelinearity r byl postupně nastaven na hodnoty $r = 0.01$, $r = 0.1$, $r = 0.25$, $r = 0.5$, $r = 0.75$ a $r = 1$.

6.1. Odolnost strategie vůči náhodným chybám pozorování veličin soustavy

Výsledky testů odolnosti strategie vůči náhodným chybám pozorování veličin soustavy jsou uvedeny na obr. 4. U nelineárních mřížek 3-D stavů bylo dosaženo se vzrůstající hodnotou chyby pozorování veličin soustavy nejnižší hodnoty integrálního kritéria kvality regulace s mřížkou (12/0.1, 8, 4; 2). Nejhuře se chovala mřížka (12/0.5, 8, 4; 2), u které integrální kritérium kvality regulace začalo prudce stoupat již při úrovni 50 % chyb pozorování veličin.

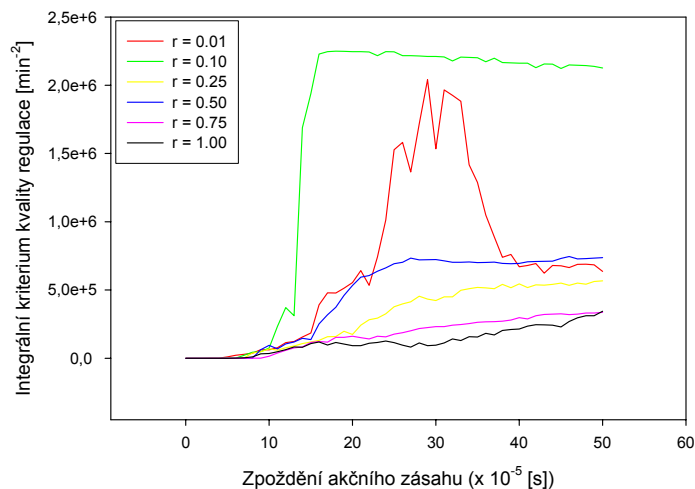


Obr. 4.: Odolnost strategie vůči náhodným chybám pozorování veličin soustavy

6.2. Odolnost strategie vůči zpoždění akčního zásahu

Na obr. 5 jsou shrnuty výsledky těchto simulací. Nad hodnotou přibližně 10×10^{-5} [s] začíná u nelineárních 3-D mřížek prudce stoupat hodnota integrálního kritéria kvality regulace. Nejlépe se chová mřížka (12/1, 8, 4; 2) u které je nejméně strmý vzestup hodnoty integrálního kritéria kvality regulace. Se vzrůstající nelinearitou se u 3-D mřížek značně zhoršuje odolnost strategie vůči zpoždění akčního zásahu.

3-D stav (12/r, 08, 04; 2) - řízení

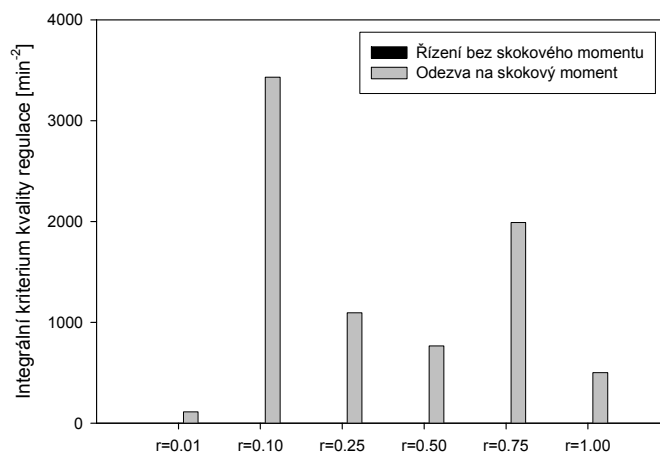


Obr. 5.: Odolnost strategie vůči zpoždění akčního zásahu

6.3. Odezva na skokový moment

Výsledky testů odezvy řízení na skokový moment jsou uvedeny na obr. 6. U všech testovaných 3-D mřížek byla hodnota integrálního kritéria kvality regulace bez skokového momentu rovna 0. Po zatížení skokovým momentem se nejlépe chovala mřížka (12/0.01, 8, 4; 2) a (12/1, 8, 4; 2). Nejhuře se chovala mřížka (12/0.1, 8, 4; 2).

3-D stav (12,08,04;02)



Obr. 6.: Odezva na skokový moment

6.4. Výběr vhodné nelineární mřížky

Z posuzovaných nelineárních mřížek 3-D stavu byla z hlediska odolnosti vůči chybám pozorování veličin soustavy, z hlediska zpoždění akčního zásahu a z hlediska odezvy řízení

na skokový moment zvolena nelineární mřížka (12/0.75, 8, 4; 2), protože ve všech pokusech vykazovala průměrně nejvyšší kvalitu regulace.

7. Závěr

Z výsledků experimentů vyplynulo, že volba mřížky má značný vliv na kvalitu a rychlost naučení. Dále vyplynulo, že nelineární mřížka může značně zlepšit kvalitu řízení asynchronního elektromotoru.

Z posuzovaných lineárních mřížek 3-D stavu byla jako nejlepší zvolena mřížka (12, 8, 4; 2), protože vykazovala dobrou odolnost vůči náhodné chybě pozorování veličin soustavy, dobrou odolnost vůči zpoždění akčního zásahu a dobrou odezvu na skokový moment.

Z posuzovaných nelineárních mřížek 3-D stavu byla z hlediska odolnosti vůči chybám pozorování veličin soustavy, z hlediska zpoždění akčního zásahu a z hlediska odezvy řízení na skokový moment zvolena nelineární mřížka (12/0.75, 8, 4; 2), protože ve všech pokusech vykazovala průměrně nejvyšší kvalitu regulace.

Poděkování:

Tato práce je podporována projektem MSM 0021630518 „Simulační modelování mechatronických soustav“.

Použitá literatura:

- [1] Březina T., Krejsa J.: The control of active magnetic bearing using two-phase Q-learning, Engineering Mechanics 2003, Svratka 2003.
- [2] Marada, T., Březina, T., Singule, V.: Q-learned Controller Performance of Asynchronous Electromotor. Mechatronics 2004, Warsaw, Poland, 2004.
- [3] Marada T., Singule V.: Využití Q-učení pro řízení asynchronního elektromotoru, Vybrané problémy elektrických strojů a pohonů, Lomnice u Tišnova 2003.