

Metody pomocných proměnných

doc. Ing. Petr Blaha, PhD.

16. prosince 2017

Komplexní inovace studijních programů a zvyšování kvality výuky na FEKT VUT v Brně
OP VK CZ.1.07/2.2.00/28.0193



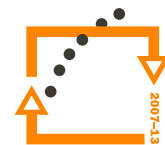
evropský
sociální
fond v ČR



EVROPSKÁ UNIE

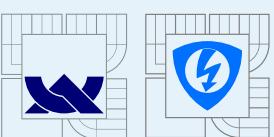


MINISTERSTVO ŠKOLSTVÍ,
MLÁDEŽE A TĚLOVÝCHOVY



OP Vzdělávání
pro konkurenceschopnost

INVESTICE DO ROZVOJE VZDĚLÁVÁNÍ



Úvod

Předpoklady

Pokračování

Analýza problému

Výsledky analýzy

Výsledky analýzy

Myšlenky metody

Odvození

Odvození-pokr

Metoda se
zpožděnými
pozorováními

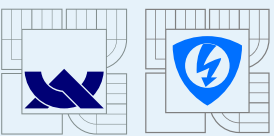
Metoda s
dodatečným
modelem

Rozšířená metoda
pomocných
proměnných

Úvod

Modelování a identifikace

Metody pomocných proměnných – strana 2 / 25



Předpoklady

Úvod

Předpoklady

Pokračování

Analýza problému

Výsledky analýzy

Výsledky analýzy

Myšlenky metody

Odvození

Odvození-pokr

Metoda se
zpožděnými
pozorováními

Metoda s
dodatečným
modelem

Rozšířená metoda
pomocných
proměnných

Uvažujme systém popsany diferenční rovnicí

$$A(q^{-1})y(k) = B(q^{-1})u(k) + \nu(k) \quad (1)$$

kde $A(q^{-1}) = 1 + a_1q^{-1} + \dots + a_{n_a}q^{-n_a}$ a

$B(q^{-1}) = b_1q^{-1} + \dots + a_{n_b}q^{-n_b}$ Výše uvedená rovnice (1) se dá přepsat ve tvaru

$$y(k) = \varphi^T(k)\theta + \nu(k) \quad (2)$$

kde $\varphi^T(k) =$

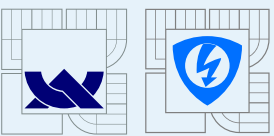
$$\begin{pmatrix} -y(k-1) & \dots & -y(k-n_a) & u(k-1) & \dots & u(k-n_b) \end{pmatrix}$$

a $\theta^T = (a_1 \quad \dots \quad a_{n_a} \quad b_1 \quad \dots \quad b_{n_b})$.

Rovnice (2) tvarem odpovídá rovnici, která se používá u lineární regrese.

Modelování a identifikace

Metody pomocných proměnných – strana 3 / 25



Pokračování

- Úvod
- Předpoklady
- Pokračování**
- Analýza problému
- Výsledky analýzy
- Výsledky analýzy
- Myšlenky metody
- Odvození
- Odvození-pokr
- Metoda se zpožděnými pozorováními
- Metoda s dodatečným modelem
- Rozšířená metoda pomocných proměnných

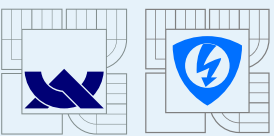
Minimalizaci kritéria

$$J_N(\theta) = \frac{1}{N} \sum_{k=1}^N \varepsilon^2(k)$$

řeší rovnice

$$\hat{\theta} = \left[\frac{1}{N} \sum_{k=1}^N \varphi(k) \varphi^T(k) \right]^{-1} \left[\frac{1}{N} \sum_{k=1}^N \varphi(k) y(k) \right]$$

Algoritmy pro výpočet odhadu zůstávají stejné. Statistické vlastnosti záleží na tom, zda je vektor měřených veličin $\varphi(k)$ apriorně známý, nebo vzniká realizací stochastického procesu.



Analýza problému

- Úvod
- Předpoklady
- Pokračování
- Analýza problému**
- Výsledky analýzy
- Výsledky analýzy
- Myšlenky metody
- Odvození
- Odvození-pokr
- Metoda se zpožděnými pozorováními
- Metoda s dodatečným modelem
- Rozšířená metoda pomocných proměnných

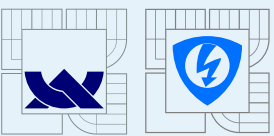
Uvažujme, že data vyhovují rovnici

$$y(k) = \varphi^T(k)\theta_0 + \nu(k)$$

kde θ_0 je vektor skutečných parametrů a vstupní signál $u(k)$ nezávisí na $\nu(k)$. Pokud je odhad $\hat{\theta}$ správný, měl by se shodovat s θ_0 .

$$\begin{aligned}\hat{\theta} - \theta_0 &= \left[\frac{1}{N} \sum_{k=1}^N \varphi(k) \varphi^T(k) \right]^{-1} \left[\frac{1}{N} \sum_{k=1}^N \varphi(k) y(k) - \right. \\ &\quad \left. - \left\{ \frac{1}{N} \sum_{k=1}^N \varphi(k) \varphi^T(k) \right\} \theta_0 \right] = \\ &= \left[\frac{1}{N} \sum_{k=1}^N \varphi(k) \varphi^T(k) \right]^{-1} \left[\frac{1}{N} \sum_{k=1}^N \varphi(k) \nu(k) \right] \stackrel{?}{=} 0\end{aligned}$$

Modelování a identifikace



Výsledky analýzy

- Úvod
- Předpoklady
- Pokračování
- Analýza problému
- Výsledky analýzy**
- Výsledky analýzy
- Myšlenky metody
- Odvození
- Odvození-pokr
- Metoda se zpožděnými pozorováními
- Metoda s dodatečným modelem
- Rozšířená metoda pomocných proměnných

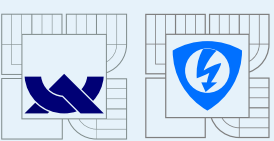
Předchozí rovnice je rovna nule, když

[1] $E\varphi(k)\varphi^T(k)$ není singulární

[2] $E\varphi(k)\nu(k) = 0$

Podmínka [1] je většinou splněna. Je zde několik vyjímek.

- vstup není perzistentně budícím signálem dostatečného řádu
- data jsou bez šumu $\nu(k) = 0$ a je zvolen model vysokého řádu (společné kořeny)
- vstup $u(k)$ je tvořen jako zpětná vazba nízkého řádu z výstupu



Výsledky analýzy

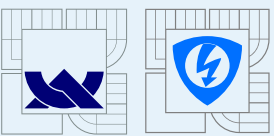
- Úvod
- Předpoklady
- Pokračování
- Analýza problému
- Výsledky analýzy
- Výsledky analýzy**
- Myšlenky metody
- Odvození
- Odvození-pokr
- Metoda se zpožděnými pozorováními
- Metoda s dodatečným modelem
- Rozšířená metoda pomocných proměnných

Podmínka **[2]** většinou splněna není. Vyjimku tvoří případ, kdy je $\nu(k)$ bílý šum. Pokud se nejedná o bílý šum, je většinou korelovaný s minulými výstupy $y(k)$ a podmínka splněna není - odhad je posunutý.

Posunutí je malé v případě vysokého poměru signál / šum. Někdy v přiměřené míře nevadí (návrh regulátoru - navrhuje se jako dostatečně robustní).

V praxi se používají dvě skupiny metod, které dávají neposunutý odhad i v případě nesplnění podmínky **[2]**

- metody pomocných proměnných
- metoda chyby predikce



Myšlenky metody

- Úvod
- Předpoklady
- Pokračování
- Analýza problému
- Výsledky analýzy
- Výsledky analýzy
- Myšlenky metody**
- Odvození
- Odvození-pokr
- Metoda se zpožděnými pozorováními
- Metoda s dodatečným modelem
- Rozšířená metoda pomocných proměnných

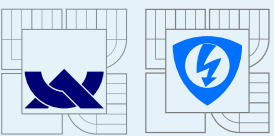
Hlavní myšlenkou metod pomocných proměnných je vytvoření nového vektoru pozorování $\varphi(k)$ tak, aby byl co nejvíce korelovaný s nezašuměnými daty ($y(k)$ a $u(k)$) a co nejméně korelovaný s šumem $\nu(k)$.

Snažíme se dosáhnout výsledku

$$E\varphi(k)\nu(k) = R_{\varphi\nu}(0) = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N \varphi(i)\nu(i) = 0$$

Používají se dva přístupy:

- metoda pomocných proměnných se zpožděnými pozorováními
- metoda pomocných proměnných s dodatečným modelem



Odvození

- Úvod
- Předpoklady
- Pokračování
- Analýza problému
- Výsledky analýzy
- Výsledky analýzy
- Myšlenky metody
- Odvození**
- Odvození-pokr
- Metoda se zpožděnými pozorováními
- Metoda s dodatečným modelem
- Rozšířená metoda pomocných proměnných

Uvažujme rovnici

$$y(k) = \varphi^T(k)\theta + \nu(k)$$

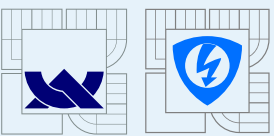
Násobením obou stran rovnice $\varphi(k)$ a sumací pro $k = 1, \dots, N$

$$\sum_{k=1}^N \varphi(k)y(k) = \left[\sum_{k=1}^N \varphi(k)\varphi^T(k) \right] \theta + \sum_{k=1}^N \varphi(k)\nu(k)$$

Pro nekorelované $\varphi(k)$ a $\nu(k)$ je poslední člen roven 0. Pro vektor θ platí

$$\theta = \left[\sum_{k=1}^N \varphi(k)\varphi^T(k) \right]^{-1} \sum_{k=1}^N \varphi(k)y(k)$$

Modelování a identifikace



Pokračování odvození

- Úvod
- Předpoklady
- Pokračování
- Analýza problému
- Výsledky analýzy
- Výsledky analýzy
- Myšlenky metody
- Odvození
- Odvození-pokr**
- Metoda se zpožděnými pozorováními
- Metoda s dodatečným modelem
- Rozšířená metoda pomocných proměnných

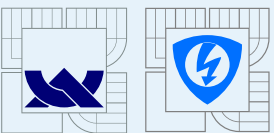
V případě, že $\varphi(k)$ a $\nu(k)$ jsou korelované, lze použít vektor $\zeta(k)$ (dimenze n), který je korelovaný s $\varphi(k)$ a nekorelovaný s $\nu(k)$. Potom

$$\sum_{k=1}^N \zeta(k)y(k) = \left[\sum_{k=1}^N \zeta(k)\varphi^T(k) \right] \theta + \sum_{k=1}^N \zeta(k)\nu(k)$$

Zde je poslední člen podle předpokladů nulový a proto

$$\theta = \left[\sum_{k=1}^N \zeta(k)\varphi^T(k) \right]^{-1} \sum_{k=1}^N \zeta(k)y(k) \quad (3)$$

Získáváme tím **neposunutý odhad**. Vektor $\zeta(k)$ se nazývá vektor **pomocných proměnných (instrumentů)**.



Úvod

Metoda se
zpožděnými
pozorováními

Motivační příklad

Motivační příklad -
pokračování

Motivační příklad -
pokračování 2

Obecný případ

Algoritmus adaptace

Shrnutí

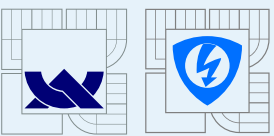
Metoda s
dodatečným
modelem

Rozšířená metoda
pomocných
proměnných

Metoda se zpožděnými pozorováními

Modelování a identifikace

Metody pomocných proměnných – strana 11 / 25



Motivační příklad

Úvod

Metoda se
zpožděnými
pozorováními

Motivační příklad

Motivační příklad -
pokračování

Motivační příklad -
pokračování 2

Obecný případ

Algoritmus adaptace

Shrnutí

Metoda s
dodatečným
modelem

Rozšířená metoda
pomocných
proměnných

Uvažujme dynamický systém prvního řádu

$$y(k+1) = -a_1 y(k) + b_1 u(k) + c_1 e(k) + e(k+1)$$

Uvažujme **apriorní odhad** výstupu

$$\hat{y}(k+1|k) = -\hat{a}_1(k)y(k) + \hat{b}_1(k)u(k) = \varphi^T(k+1)\hat{\theta}(k)$$

kde

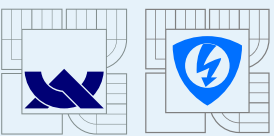
$$\hat{\theta}(k) = \begin{pmatrix} \hat{a}_1(k) \\ \hat{b}_1(k) \end{pmatrix} \quad \varphi(k+1) = \begin{pmatrix} -y(k) \\ u(k) \end{pmatrix}$$

a **aposteriorní odhad** výstupu

$$\hat{y}(k+1|k+1) = \varphi^T(k+1)\hat{\theta}(k+1)$$

Modelování a identifikace

Metody pomocných proměnných – strana 12 / 25



Motivační příklad - pokračování

Úvod

Metoda se
zpožděnými
pozorováními

Motivační příklad

Motivační příklad -
pokračování

Motivační příklad -
pokračování 2

Obecný případ

Algoritmus adaptace

Shrnutí

Metoda s
dodatečným
modelem

Rozšířená metoda
pomocných
proměnných

Chyba aposteriorního odhadu

$$\begin{aligned}\varepsilon(k+1|k+1) &= y(k+1) - \varphi^T(k+1)\hat{\theta}(k+1) = \\ &= \varphi^T(k+1) \left[\theta_0 - \hat{\theta}(k+1) \right] + c_1 e(k) + e(k+1)\end{aligned}$$

kde

$$\hat{\theta}(k+1) = \begin{pmatrix} \hat{a}_1(k+1) \\ \hat{b}_1(k+1) \end{pmatrix} \quad \theta_0 = \begin{pmatrix} a_1 \\ b_1 \end{pmatrix}$$

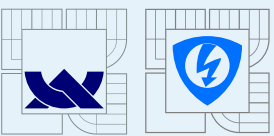
Uvažujme, že se odhad shoduje se skutečnými parametry

$$\hat{\theta}(k+1) = \theta_0$$

$$E\varphi(k+1)\varepsilon(k+1) = \begin{pmatrix} -Ey(k)\varepsilon(k+1) \\ 0 \end{pmatrix} = \begin{pmatrix} c_1 Ee^2(k) \\ 0 \end{pmatrix} = \begin{pmatrix} c_1 \sigma^2 \\ 0 \end{pmatrix}$$

Modelování a identifikace

Metody pomocných proměnných – strana 13 / 25



Motivační příklad - pokračování 2

Úvod

Metoda se
zpožděnými
pozorováními

Motivační příklad

Motivační příklad -
pokračování

Motivační příklad -
pokračování 2

Obecný případ

Algoritmus adaptace

Shrnutí

Metoda s
dodatečným
modelem

Rozšířená metoda
pomocných
proměnných

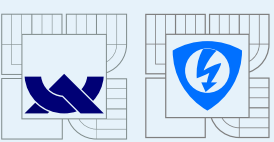
Pokusme se nyní nahradit první člen ve vektoru pozorování $\varphi(k+1)$ jeho zpožděnou hodnotou o jeden krok. Tím vytvoříme vektor pomocných proměnných.

$$\zeta_d(k+1) = \begin{pmatrix} -y(k-1) \\ u(k) \end{pmatrix}$$

Protože $y(k-1)$ je závislý na členech $e(k-1)$ a $e(k-2)$, zřejmě platí

$$-Ey(k-1)\varepsilon(k+1) = 0 \rightarrow E\zeta_d(k+1)\varepsilon(k+1) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

Je vidět, že touto jednoduchou úpravou dostaneme **neposunutý odhad**.



Obecný případ

Úvod

Metoda se
zpožděnými
pozorováními

Motivační příklad

Motivační příklad -
pokračování

Motivační příklad -
pokračování 2

Obecný případ

Algoritmus adaptace

Shrnutí

Metoda s
dodatečným
modelem

Rozšířená metoda
pomocných
proměnných

Předpokládaný obecný model je ve tvaru

$$A(q^{-1})y(k) = q^{-d}B(q^{-1})u(k) + C(q^{-1})e(k)$$

Vektor pomocných proměnných je dán tvarem

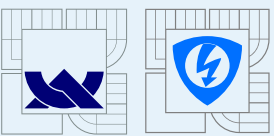
$$\zeta_d(k+1) = \begin{pmatrix} -y(k - n_c) \\ -y(k - n_c - 1) \\ \vdots \\ -y(k - n_c - n_a + 1) \\ u(k) \\ u(k - 1) \\ \vdots \\ u(k - n_b + 1) \end{pmatrix}$$

Aby byl získaný odhad neposunutý, musí zpoždění
pozorování n_c musí splňovat podmínku

$$n_c \geq \deg C(q^{-1})$$

Modelování a identifikace

Metody pomocných proměnných – strana 15 / 25



Algoritmus adaptace

Úvod

Metoda se
zpožděnými
pozorováními

Motivační příklad
Motivační příklad -
pokračování
Motivační příklad -
pokračování 2
Obecný případ

Algoritmus adaptace

Shrnutí

Metoda s
dodatečným
modelem

Rozšířená metoda
pomocných
proměnných

Off-line odhad popisuje rovnice (3). On-line metoda zůstává stejná jako u metody nejmenších čtverců.

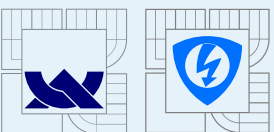
$$\hat{\theta}(k) = \hat{\theta}(k-1) + K(k)\varepsilon(k)$$

$$\varepsilon(k) = y(k) - \varphi^T(k)\hat{\theta}(k-1)$$

$$K(k) = P(k)\zeta(k) = \frac{P(k-1)\zeta(k)}{\lambda + \varphi^T(k)P(k-1)\zeta(k)}$$

$$P(k) = \frac{1}{\lambda} \left[P(k-1) - \frac{P(k-1)\zeta(k)\varphi^T(k)P(k-1)}{\lambda + \varphi^T(k)P(k-1)\zeta(k)} \right]$$

Je vidět, že se v rovnicích pro klasickou MNČ nahradil vektor $\varphi(k)$ vektorem $\zeta(k)$ a vektor $\varphi^T(k)$ **zůstal zachován**. Variantu bez exponenciálního zapomínání získáme tak, že dosadíme $\lambda = 1$.



Shrnutí

Úvod

Metoda se
zpožděnými
pozorováními

Motivační příklad
Motivační příklad -
pokračování
Motivační příklad -
pokračování 2

Obecný případ

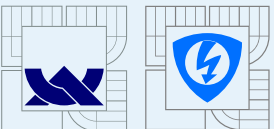
Algoritmus adaptace

Shrnutí

Metoda s
dodatečným
modelem

Rozšířená metoda
pomocných
proměnných

- Tuto metodu se doporučuje inicializovat. Používá se většinou klasická rekurzivní metoda nejmenších čtverců. V praxi se volí 5-8 násobek neznámých parametrů.
- Funguje správně v případě vysokofrekvenčního šumu (šum měření).
- Často vyžaduje zvýšení vzorkovací frekvence.
- Nepotřebuje provádět identifikaci modelu šumu.



Úvod

Metoda se
zpožděnými
pozorováními

Metoda s
dodatečným
modelem

IVM s dodatečným
modelem

Příklad

Příklad -
pokračování

Příklad -
pokračování 2

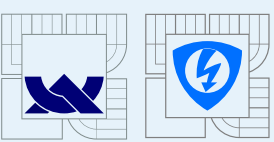
Shrnutí

Rozšířená metoda
pomocných
proměnných

Metoda s dodatečným modelem

Modelování a identifikace

Metody pomocných proměnných – strana 18 / 25



IVM s dodatečným modelem

Úvod

Metoda se
zpožděnými
pozorováními

Metoda s
dodatečným
modelem

IVM s dodatečným
modelem

Příklad

Příklad -
pokračování

Příklad -
pokračování 2

Shrnutí

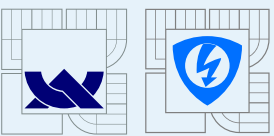
Rozšířená metoda
pomocných
proměnných

Cíl je stejný, jako u metody IVMd se zpožděnými pozorováními. V tomto případě uvažujeme strukturu modelu

$$A(q^{-1})y(k) = q^{-d}B(q^{-1})u(k) + A(q^{-1})e(k)$$

Tento model odpovídá situaci, kdy šum $e(k)$ vstupuje na výstupu soustavy s přenosem $F(z^{-1}) = \frac{z^{-d}B(z^{-1})}{A(z^{-1})}$. Opět předpokládáme, že je porucha $e(k)$ nezávislá na vstupním signálu $u(k)$ a má nulovou střední hodnotu a nenulový rozptyl σ^2 .

Cílem je určit neposunuté odhady koeficientů přenosu, aniž bychom museli uvažovat modelování poruchy.



Příklad

Úvod

Metoda se
zpožděnými
pozorováními

Metoda s
dodatečným
modelem
IVM s dodatečným
modelem

Příklad

Příklad -
pokračování

Příklad -
pokračování 2

Shrnutí

Rozšířená metoda
pomocných
proměnných

Uvažujme systém popsany rovnicí

$$y(k+1) = -a_1 y(k) + b_1 u(k) + a_1 e(k) + e(k+1)$$

Dále uvažujme prediktor ve tvaru

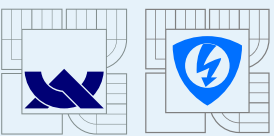
$$\hat{y}(k+1|k) = -\hat{a}_1(k)y(k) + \hat{b}_1(k)u(k) = \varphi^T(k+1)\hat{\theta}(k)$$

kde

$$\hat{\theta}(k) = \begin{pmatrix} \hat{a}_1(k) \\ \hat{b}_1(k) \end{pmatrix} \quad \varphi(k+1) = \begin{pmatrix} -y(k) \\ u(k) \end{pmatrix}$$

a a posteriori odhad výstupu

$$\hat{y}(k+1|k+1) = \varphi^T(k+1)\hat{\theta}(k+1)$$



Příklad - pokračování

Úvod

Metoda se
zpožděnými
pozorováními

Metoda s
dodatečným
modelem
IVM s dodatečným
modelem

Příklad

**Příklad -
pokračování**

Příklad -
pokračování 2

Shrnutí

Rozšířená metoda
pomocných
proměnných

Chyby predikce (apriorní a aposteriorní) jsou dány vztahy

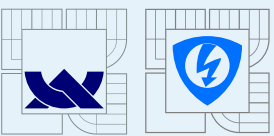
$$\varepsilon(k+1|k) = y(k+1) - \hat{y}(k+1|k)$$

$$\varepsilon(k+1|k+1) = y(k+1) - \hat{y}(k+1|k+1)$$

Definujme dodatečný model predikce, který bude tvořit pomocnou proměnnou

$$y_{IVM}(k) = -\hat{a}_1 y_{IVM}(k-1) + \hat{b}_1 u(k-1)$$

Tato rovnice nezávisí na minulé hodnotě výstupu, ale na jeho odhadu. $y(k-1)$ je zde nahrazeno $y_{IVM}(k-1)$. Tato nová proměnná bude méně ovlivněná působící poruchou, což povede k získání neposunutého odhadu parametrů.



Příklad - pokračování 2

Úvod

Metoda se
zpožděnými
pozorováními

Metoda s
dodatečným
modelem

IVM s dodatečným
modelem

Příklad

Příklad -
pokračování

Příklad -
pokračování 2

Shrnutí

Rozšířená metoda
pomocných
proměnných

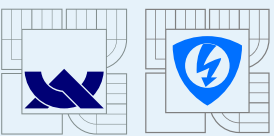
Vektor pomocných proměnných je definován jako

$$\zeta_m(k+1) = \begin{pmatrix} -y_{IVM}(k) \\ u(k) \end{pmatrix}$$

V obecném případě bude situace obdobná.

Pomocné proměnné $y_{IVM}(k)$ jsou získány z dodatečného modelu

$$\hat{A}(k, q^{-1})y_{IVM}(k) = q^{-d}\hat{B}(k, q^{-1})u(k)$$



Shrnutí

Úvod

Metoda se
zpožděnými
pozorováními

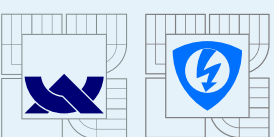
Metoda s
dodatečným
modelem
IVM s dodatečným
modelem

Příklad
Příklad -
pokračování
Příklad -
pokračování 2

Shrnutí

Rozšířená metoda
pomocných
proměnných

- Algoritmus adaptace zůstává stejný jako v případě rekurzivní metody nejmenších čtverců, pouze se změněným vektorem pozorování $\varphi(k)$.
- Abychom mohli spočítat pomocné proměnné, musíme znát dopředu přibližný odhad koeficientů přenosu. Z tohoto důvodu musí být metoda opět inicializována rekurzivní metodou nejmenších čtverců. Horizont inicializace se volí stejně jako v předchozím případě, t.j. 5-8 krát počet neznámých koeficientů.
- Jestliže je horizont krátký a získaný odhad není dostatečně přesný, může popsáný algoritmus divergovat.



Úvod

Metoda se
zpožděnými
pozorováními

Metoda s
dodatečným
modelem

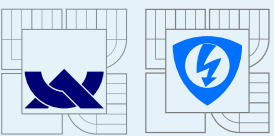
Rozšířená metoda
pomocných
proměnných

Rozšířená metoda
pomocných
proměnných

Rozšířená metoda pomocných proměnných

Modelování a identifikace

Metody pomocných proměnných – strana 24 / 25



Rozšířená metoda pomocných proměnných

Úvod

Metoda se
zpožděnými
pozorováními

Metoda s
dodatečným
modelem

Rozšířená metoda
pomocných
proměnných

Rozšířená metoda
pomocných
proměnných

Jedná se o zobecnění metody IVM o $\dim \zeta > n$ a případnou předfiltraci tohoto vektoru filtrem $F(q^{-1})$

$$\hat{\theta} = \arg \min_{\theta} \left\| \left[\sum_{k=1}^N \zeta(k) F(q^{-1}) \varphi^T(k) \right] \theta - \left[\sum_{k=1}^N \zeta(k) F(q^{-1}) y(k) \right] \right\|_Q^2$$

kde $\|x\|_Q^2 = x^T Q x$ s maticí Q , která je váhovou pozitivně definitní maticí. Odhad je neposunutý, pokud

- $E \zeta(k) F(q^{-1}) \varphi^T(k)$ není singulární
- $E \zeta(k) F(q^{-1}) e(k) = 0$