

VOLD-KALMANOVA FILTRACE V PROSTŘEDÍ MATLABU

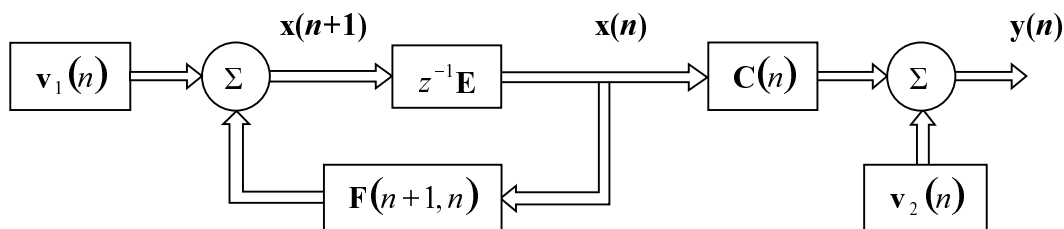
Prof. Ing. Jiří Tůma, CSc.

VŠB – Technická univerzita Ostrava, Ostrava, Česká republika

1 Úvod

Výstupem Kalmanova filtru je odhad časového vývoje stavu lineárního systému, který je dostupný jen prostřednictvím nepřímého pozorování časového vývoje jistého vektoru dat. Algoritmus odhadování je rekurzivní, tedy zvláště vhodný pro číslicové zpracování naměřených dat v reálném čase, a chyba odhadu je minimální ve smyslu statistickém, což znamená, že střední hodnota druhé mocniny chyby je minimální. Bez bližšího popisu podstaty rekurzivních výpočtů bude v úvodu popsáno jen zadání úlohy filtrace jako referenční metody k Vold-Kalmanově filtraci. Označení veličin bude převzato z knihy [Haykin 1996].

Základem filtru jsou dvě rovnice. První rovnice (process equation) popisuje vývoj stavu procesu popsaného v časovém okamžiku n vektorem stavových proměnných $\mathbf{x}(n)$ tak, že je definována souvislost mezi stavovým vektorem v okamžiku n a v následujícím okamžiku $n+1$. Stav procesu v okamžiku $n+1$ je ovlivněn náhodným vektorem $\mathbf{v}_1(n)$. Složky tohoto náhodného vektoru jsou členy bílé posloupnosti, tj. lze je označit za nekorelovaný šum. Druhá rovnice (measurement equation) modeluje měření stavu. Zatímco první rovnice byla v principu dynamická – diferenční, představuje druhá rovnice prostou transformaci hodnot vektoru stavu prostřednictvím transformační matice na hodnoty přístupné měření $\mathbf{y}(n)$. Transformovaný stav ovlivněn aditivním šumem $\mathbf{v}_2(n)$. Vektor šumu obsahuje také složky představující bílé posloupnosti. Blokové schéma vzájemných vazeb lze znázornit na obr. 1. Symbol z^{-1} znázorňuje posunutí o jeden vzorek (\mathbf{E} je jednotková matice).



Obr. 1 Signálový diagram Kalmanova filtru

Výstupem Kalmanova filtru jsou odhady vektoru stavu procesu $\hat{\mathbf{x}}(n|\mathbf{y}(1), \dots, \mathbf{y}(n))$ v časovém okamžiku n na základě měřených hodnot $\mathbf{y}(1), \dots, \mathbf{y}(n)$ a odhadu počátečního stavu $\hat{\mathbf{x}}(1|\mathbf{y}(1))$ a korelačního vektoru chyby tohoto počátečního odhadu.

Zadání úlohy Kalmanovy filtrace lze aplikovat na problém extrakce harmonické složky z měřeného signálu. Stav procesu lze v této úloze považovat za hledanou harmonickou složku, která v součtu se zbývajícími složkami signálu představuje měřený signál. Popis vývoje stavu procesu představující tuto harmonickou složku s případně měnící se frekvencí bude předvedeno v další části referátu. Modelování měření představuje sumaci harmonické složky se zbytkem složek obsažených v signálu. Jistý problém však představují budící signály $\mathbf{v}_1(n)$ a $\mathbf{v}_2(n)$ signálového diagramu z obr. 1. Jejich rozptyly nejsou předem známy a navíc signál $\mathbf{v}_2(n)$ nemusí být dominující bílý šum, ale další harmonická složka o jiné frekvenci.

Publikované práce Volda a Leuridana počínaje rokem 1993 se přesné aplikaci metodiky Kalmanova filtrace vyhýbají a těží jen z postupu výpočtu, jehož detaily nejsou publikovány.

V úvahách Volda se respektuje neznalost rozptylů obou zmíněných budících posloupností a ve výpočtu se ovládá jen jejich vzájemný poměr, který ovlivňuje frekvenční vlastnosti filtru. I když lze pochybovat o významu dílčího označení tohoto způsobu filtrace jménem Kalman, v principu je myšlenka filtrace navržena Voldem a Leuridanem inspirující a otvírá nové možnosti řádové analýzy signálů se známou frekvencí, která je odvozena z měření otáček. Popularizace této metody filtrace byla zajištěna její implementací do signálového analyzátoru PULSE firmy Brüel & Kjær. Jak je ukázáno v tomto referátu, rovněž prostředí Matlabu umožňuje tuto filtraci výhodně provádět.

2 Vold-Kalmanův filtr první generace

Účelem první generace filtru tohoto typu bylo extrahování modulované harmonické složky o známé frekvenci pro analýzu časového průběhu změn její amplitudy například za rozběhu stroje. Zvýšená amplituda může indikovat například rezonanci. Vold-Kalmanův filtr může být použit také k odfiltrování zvolené části spektra pro systémy analyzující kvalitu zvuku.

2.1 Datová rovnice

Jestliže z měřeného signálu se vzorky $y(n)$ je žádoucí extrahovat jednu harmonickou složku, pak datová rovnice vyjadřuje prostý fakt, že měřený signál je složen z extrahované harmonické složky se vzorky označenými $x(n)$ a zbytku signálu se vzorky označenými $\eta(n)$, kde $n = 1, \dots, N$ znamená pořadí vzorku. Její tvar pro vzorky s pořadím n je následující

$$y(n) = x(n) + \eta(n). \quad (1)$$

Harmonická složka může být amplitudově nebo fázově modulována. Zbytkové složky signálu mohou představovat nejen širokopásmový šum, ale také další harmonické složky.

Obecný tvar datové rovnice pro vzorky s pořadím n a s počtem P sledovaných harmonických složek $x_i(n)$ je

$$y(n) = \sum_{i=1}^P x_i(n) + \eta(n). \quad (2)$$

Tento vztah platí pro každý naměřený a extrahovaný vzorek signálu.

2.2 Strukturální rovnice pro harmonický signál

V čase spojitý harmonický signál $x(t) = A \cos(\omega t)$ je řešením diferenciální rovnice druhého řádu s nulovým součinitelem tlumení. Pro případ vzorkování signálu v diskrétních časových okamžicích $t_n = n \Delta t$, $n = 0, 1, 2, \dots$ je harmonický signál řešením diferenční rovnice rovněž druhého řádu, jejíž charakteristická rovnice má dva komplexně sdružené kořeny $z_1 = \exp(j\omega \Delta t)$ a $z_2 = \exp(-j\omega \Delta t)$, kde ω je úhlová frekvence. Obecné řešení diferenční rovnice druhého řádu rovnice se zapisuje ve tvaru $x(n) = C(z_1^n + z_2^n)$. Charakteristický polynom tvoří součin kořenových činitelů $(z - z_1)(z - z_2)$, ze kterých lze rekonstruovat výchozí homogenní diferenční rovnici

$$x(n) - 2 \cos(\omega \Delta t)x(n-1) + x(n-2) = 0, \quad (3)$$

přičemž její koeficient u $x(n-1)$ je vhodné označit $c(n) = 2 \cos(\omega \Delta t)$.

K výpočtu posloupností vzorků postačují hodnoty prvních dvou vzorků a samozřejmě velikost úhlové frekvence. Výsledkem řešení rovnice (3) s konstantní velikostí ω je harmonický signál o konstantní amplitudě.

Harmonické složky měřených signálů z rotačních strojů mění ve skutečnosti svou frekvenci a amplitudu, tj. stávají se modulovanými harmonickými signály. Jednu harmonickou složku může popisovat rovnice

$$x(n) - 2 \cos(\omega \Delta t)x(n-1) + x(n-2) = \varepsilon(n), \quad (4)$$

kde $\varepsilon(n)$ je budící funkce, která spolu s velikostí okamžité úhlové frekvence ω určuje řešení diferenční rovnice (4).

Přenos budící funkce $\varepsilon(n)$ na posloupnost vzorků $x(n)$ není stabilní, protože kořeny charakteristické rovnice leží na jednotkové kružnici. Zesílení je nekonečné pro složku o úhlové frekvenci ω .

2.3 Soustava datových a strukturálních rovnic

Datové rovnice pro všechny naměřené vzorky signálu a neznámé harmonické složky o celkovém počtu N s individuálními vzorky lze zapsat takto:

$$\begin{bmatrix} y(1) \\ y(2) \\ \dots \\ y(N) \end{bmatrix} - \begin{bmatrix} x(1) \\ x(2) \\ \dots \\ x(N) \end{bmatrix} = \begin{bmatrix} \eta(1) \\ \eta(2) \\ \dots \\ \eta(N) \end{bmatrix} \quad (5)$$

Pohodlnější je používání zkráceného zápisu pomocí vektorů

$$\mathbf{y} - \mathbf{x} = \boldsymbol{\eta}. \quad (6)$$

Velikost zbytkového signálu jako vektoru lze ohodnotit jeho euklidovskou normou (zobecněnou délkou), která v případě reálných signálů představuje součet čtverců jednotlivých vzorků $\eta(n) = y(n) - x(n)$, $n = 1, \dots, N$. Vektorový zápis této euklidovské normy zjednodušuje transpozice vektoru $\boldsymbol{\eta}^T$

$$\boldsymbol{\eta}^T \boldsymbol{\eta} = (\mathbf{y}^T - \mathbf{x}^T)(\mathbf{y} - \mathbf{x}). \quad (7)$$

Podobně jako datové rovnice lze také strukturální rovnice v počtu $N-2$ pro jednotlivé vzorky uspořádat následujícím způsobem

$$\begin{bmatrix} 1 & -c(1) & 1 & 0 & \dots & 0 & 0 & 0 \\ 0 & 1 & -c(2) & 1 & \dots & 0 & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & 0 & \dots & 1 & -c(N-2) & 1 \end{bmatrix} \begin{bmatrix} x(1) \\ x(2) \\ \dots \\ x(N) \end{bmatrix} = \begin{bmatrix} \varepsilon(3) \\ \varepsilon(4) \\ \dots \\ \varepsilon(N) \end{bmatrix}. \quad (8)$$

Menší počet rovnic je dán řádem diferenční rovnice. Jako v případě datových rovnic je pohodlnější pracovat s maticovým zápisem soustavy strukturálních rovnic.

$$\mathbf{A} \mathbf{x} = \boldsymbol{\varepsilon}, \quad (9)$$

kde \mathbf{A} je pásová třídiagonální obdélníková matice s N sloupci a $N-2$ řádky.

Podobně jako zbytkový signál $\eta(n)$ lze také velikost budící funkce se vzorky $\varepsilon(n)$ $n = 1, \dots, N$ ohodnotit její euklidovskou normou. Transpozice vektoru zjednodušuje zápis součtu čtverců:

$$\boldsymbol{\varepsilon}^T \boldsymbol{\varepsilon} = \mathbf{x}^T \mathbf{A}^T \mathbf{A} \mathbf{x}. \quad (10)$$

Násobením matice \mathbf{A}^T (horní index T označuje transpozici) maticí \mathbf{A} vznikne čtvercová symetrická matice o rozměru $N \times N$. Euklidovská norma vektoru obsahujícího budící funkce

má tvar kvadratická formy $\mathbf{x}^T \mathbf{A}^T \mathbf{A} \mathbf{x}$ v proměnné \mathbf{x} , o které lze podle základních vět teorie matic uvést, že je nezáporná, tj. $\mathbf{x}^T \mathbf{A}^T \mathbf{A} \mathbf{x} \geq 0$, pro libovolný nenulový vektor budících funkcí $\mathbf{x} \neq \mathbf{0}$. Tato vlastnost se nazývá pozitivní semidefinitnost čtvercové symetrické matice $\mathbf{A}^T \mathbf{A}$. Semidefinitnost matice lze zdůvodnit tím, že tato kvadratická forma může teoreticky nabýt také nulové hodnoty pro nenulový vektor \mathbf{x} .

Počet nenulových diagonál u matice $\mathbf{A}^T \mathbf{A}$ se zvýší na 5, přičemž nenulové prvky jsou na dvou diagonálách pod a nad hlavní diagonálou této čtvercové pásové matice.

2.4 Globální řešení datových a strukturálních rovnic

Výsledkem filtrace má být nalezení neznámého vektoru \mathbf{x} , který musí vyhovovat soustavě datových a strukturálních rovnic. Počet prvků vektoru \mathbf{x} je N a počet rovnic $N + (N - 2) = 2N - 2$. Mezi neznámé patří také vektory $\boldsymbol{\varepsilon}$ ($N-2$ prvků) a $\boldsymbol{\eta}$ (N prvků), pak počet neznámých $3N - 2$ je vyšší než počet rovnic, tj. soustava rovnic je vzhledem ke zmíněným třem neznámým vektorům nedourčena. K soustavě rovnic mohou být připojeny podmínky, podle kterých je požadováno, aby vektory $\boldsymbol{\varepsilon}, \boldsymbol{\eta}$ měly minimální euklidovskou normu, a aby tyto normy byly v žádoucím vztahu, tj. aby bylo minimalizováno kritérium ve tvaru

$$J = r^2 \boldsymbol{\varepsilon}^T \boldsymbol{\varepsilon} + \boldsymbol{\eta}^T \boldsymbol{\eta} = \mathbf{x}^T (r^2 \mathbf{A}^T \mathbf{A} + \mathbf{E}) \mathbf{x} - \mathbf{y}^T \mathbf{x} - \mathbf{x}^T \mathbf{y}. \quad (11)$$

kde r je váhový koeficient, který určuje předepsaný vzájemný vztah mezi euklidovskými normami zbytkového signálu a budící funkce. Jak bude dále ukázáno, podmínka minimální velikosti váženého součtu a zvolený váhový koeficient dávají vypočtenému vektoru výhodné vlastnosti, protože realizují filtr s výhodným průběhem frekvenční charakteristiky představující pásmovou propust. Hodnoty váhového koeficientu blízké nule zmenšují vlivnost strukturální rovnice a filtr ztrácí účinnost.

Podle monografie o speciálních maticích [Fiedler 1981] lze dokázat, že jestliže je k hlavní diagonále pozitivně definitní matice přičteno libovolně malé číslo, pak je výsledná matice pozitivně definitní, tj. pro libovolný nenulový vektor budících funkcí $\mathbf{x} \neq \mathbf{0}$ platí ostrá nerovnost $\mathbf{x}^T (r^2 \mathbf{A}^T \mathbf{A} + \mathbf{E}) \mathbf{x} > 0$. Protože je symetrická čtvercová matice $(r^2 \mathbf{A}^T \mathbf{A} + \mathbf{E})$ pozitivně definitní, je také regulární a tedy existuje její inverzní matice.

První derivace skalárního kritéria J podle vektoru \mathbf{x} se pro minimalizaci tohoto kritéria položí rovna nulovému vektoru.

$$\frac{\partial J}{\partial \mathbf{x}} = 2r^2 \mathbf{A}^T \mathbf{A} \mathbf{x} + 2(\mathbf{x} - \mathbf{y}) = \mathbf{0}, \quad (12)$$

což lze přepsat do tvaru

$$\mathbf{B} \mathbf{x} = \mathbf{y}, \quad (13)$$

kde $\mathbf{B} = r^2 \mathbf{A}^T \mathbf{A} + \mathbf{E}$. Řešení soustavy rovnic (12) je dáno následujícím vzorcem

$$\mathbf{x} = \mathbf{B}^{-1} \mathbf{y}. \quad (14)$$

Teoretický zápis řešení bude sloužit k programování algoritmu. Jednou možností je naprogramovat výpočty v Matlabu nebo sestavit speciální program, například v jazyku C nebo Visual Basicu. Klíčovým problémem jsou operace s výchozí maticí \mathbf{A} , popřípadě také s maticí \mathbf{B} . Teoreticky délka vektoru \mathbf{x} může dosahovat desítek tisíc vzorků a je proto nemožné uchovávat matici \mathbf{B} v paměti počítače celou. Prostředky Matlabu mají k dispozici úsporné ukládání řídkých (*sparse*) matic, kdy je paměť obsazena pouze nenulovými prvky. Diagonální řídkou maticí lze vytvořit příkazem `spdiags`. Tento příkaz přetransformuje čtvercovou matici tak, že její sloupce se stanou diagonálami. Dalším speciálním příkazem je

speye pro vytvoření jednotkové matice, která je uložena v paměti jako řídka. Vzájemné operace mezi řídkými maticemi tuto vlastnost zachovávají. Filtraci realizuje například funkce MyVoldKalman1:

Program pro Vold-Kalmanův filtr č.1

```
function x=MyVoldKalman1(y,dt,f,r)
% Funkce x= MyVoldKalman1 (y,dt,f,r) pro vypocet
x=inv(r^2*A'*A+E)*y
% Vstupy:
% y   vektor vstupnich dat
% f   vektor frekvence f v Hz (y a f maji stejny pocet prvku)
% dt  vzorkovaci perioda v sec, r   vahovy faktor
% Vystup:
% x   vektor minimalizuje funkci J=r*r*x'*A'*A*x+(y'-x')*(y-x)

c = 2*cos(2*pi*f*dt);
N = max(size(y));
N2 = N-2;
e = ones(N2,1);
A = spdiags([e -2*c(1:N2) e],0:2,N2,N);
AA = r*r*A'*A +speye(N,N);
x = AA\y;
```

K demonstraci obsazení paměti počítače bude uveden příklad pro vektor vstupních dat o délce 1000 prvků. Plná matice A má rozměr 998 řádků x 1000 sloupců. Jestliže pro tuto matici bude vypočten součin $A^T A$, pak obsazení paměti je následující:

```
>> whos
Name          Size          Bytes   Class
A             998x1000      7984000 double array
AA            1000x1000     8000000 double array
```

Každý prvek matice zaplní 8 bajtů. V době neustálého zvětšování paměti počítačů se paměť nemusí jevit jako limitující, ovšem zbytečné plnění paměti nulami zpomalí podstatně výpočet. Jestliže se matice deklarují jako řídké, pak obsazení paměti je následující:

```
>> whos
Name          Size          Bytes   Class
A             998x1000      39932   sparse array
AA            1000x1000     84836   sparse array
```

Úspora paměťového místa a z toho plynoucí podstatná redukce počtu operací při násobení nebo sčítání matic vede ke zrychlení výpočtů.

3 Vold-Kalmanův filtr druhé generace

Výsledek filtrace spočívající v časovém průběhu frekvenčně a amplitudově modulovaného harmonického signálu je pro technické aplikace zbytečně podrobný. Okamžitá frekvence je součástí zadání filtrace. Zájem se soustřeďuje proto jen na amplitudu sledované složky signálu představující obálku signálu. Takto definovaná úloha filtrace patří Vold-Kalmanovu filtru druhé generace.

3.1 Datová rovnice pro obálku harmonického signálu

Amplitudu lze vypočítat z amplitudy sinové složky a amplitudy kosinové složky. Tento postup užívá zbytečně mezivýpočet dílčích údajů a komplikuje řešení. Výhodnější je modelovat sledovaný signál jako součin komplexní amplitudy a komplexního harmonického signálu o jednotkové absolutní hodnotě. Datová rovnice pro jednu složku má tedy tvar

$$y(n) = x(n)\exp(j\Theta(n)) + \eta(n), \quad (15)$$

kde $x(n)$ je komplexní amplituda nebo také obálka signálu, j je imaginární jednotka a $\Theta(n)$ je průběžná fáze signálu od počátku záznamu. Podobně jako v datové rovnici (2) je $y(n)$ měřený signál a $\eta(n)$ zbytkový signál.

Průběžná fáze signálu je obecně rovna časovému integrálu úhlové frekvence. V případě vzorkovaných signálů se vypočte podle vzorce

$$\Theta(n) = \sum_{i=0}^n \omega(i)\Delta t, \quad (16)$$

kde $\omega(n)$ je okamžitá úhlová frekvence a $\Delta t = 1/f_s$ je interval vzorkování.

Soustava datových rovnic pro všechny naměřené vzorky má vektorový tvar

$$\mathbf{y} - \mathbf{C}\mathbf{x} = \boldsymbol{\eta}, \quad (17)$$

kde $\mathbf{C} = \mathbf{C}^T$ je diagonální komplexní matice tvaru

$$\mathbf{C} = \text{diag}\{\exp(j\Theta(1)), \exp(j\Theta(2)), \dots, \exp(j\Theta(N))\}. \quad (18)$$

Tato matice stejně jako vektory $\mathbf{x}, \boldsymbol{\eta}$ jsou komplexní. Charakteristika velikosti zbytkového signálu $\boldsymbol{\eta}$ se určí jako součin transponovaného komplexně sdruženého vektoru $\boldsymbol{\eta}^H$ a původního vektoru $\boldsymbol{\eta}$, který je reálný. Operace transponování matice a obrácení znaménka imaginární části jejich prvků se označuje horním indexem H . Reálné vektory budou však označeny jen znaménkem transponování. Pro euklidovskou normu vektoru zbytkového signálu platí

$$\boldsymbol{\eta}^H \boldsymbol{\eta} = (\mathbf{y}^T - \mathbf{x}^H \mathbf{C}^H)(\mathbf{y} - \mathbf{C}\mathbf{x}). \quad (19)$$

3.2 Strukturální rovnice pro obálku harmonického signálu

Strukturální rovnice filtruje posloupnost hodnot komplexní amplitudy. Omezení rychlosti jejich změn se zadává nejvyšším stupněm polynomu k , který lze skupinou vzorků obálky o počtu $k + 1$ proložit. Polynom prvního stupně je řešením diferenční rovnice druhého řádu

$$\nabla^2 x(n) = x(n) - 2x(n-1) + x(n-2) = 0. \quad (20)$$

Charakteristická rovnice této diferenční rovnice má dvojný kořen roven jednotce, což vede na zmíněný výsledek řešení.

Polynom obecného stupně k splňuje podmínku nulové difference řádu $k + 1$, tj. $\nabla^{k+1}x(n) = 0$. Protože se obálka signálu postupně mění, je třeba strukturální rovnici budít budící funkcí $\varepsilon(n)$. Konečný obecný tvar strukturální rovnice je

$$\nabla^{k+1}x(n) = \varepsilon(n), \quad (21)$$

kde k představuje řád polynomu, který se spojuje s řádem filtru změn komplexní amplitudy.

Vold-Kalmanova filtrace druhé generace užívá filtr prvního nebo druhého, popř. až třetího řádu. Strukturální rovnice pro tyto filtry mají tvar.

$$\text{První řád:} \quad x(n) - 2x(n-1) + x(n-2) = \varepsilon(n) \quad (22)$$

$$\text{Druhý řád} \quad x(n) - 3x(n-1) + 3x(n-2) - x(n-3) = \varepsilon(n) \quad (23)$$

Soustava strukturálních rovnic s váhovými koeficienty, které ovlivňují každý vzorek budící funkce podle okamžité relativní frekvence naladění, je shodná se vzorcem (8). Euklidovská norma charakterizující velikost vektoru budící funkce se však změní s ohledem na skutečnost, že vektor obálky obsahuje komplexní čísla. K transpozici je třeba přidat také změnu znaménka imaginární části, aby výsledek násobení vektorů bylo reálné číslo, tj.

$$\boldsymbol{\varepsilon}^H \boldsymbol{\varepsilon} = r^2 \mathbf{x}^H \mathbf{A}^T \mathbf{A} \mathbf{x}. \quad (24)$$

Matrice \mathbf{A} obsahuje jen konstantní prvky v závislosti na řádu filtru. Její rozměr pro filtr prvního řádu je $N \times N-2$ a pro filtr druhého řádu $N \times N-3$. Například pro filtr prvního řádu se jedná o třídiagonální obdélníkovou matici následujícího tvaru

$$\mathbf{A} = \begin{pmatrix} 1 & -2 & 1 & & & \\ & 1 & -2 & 1 & & \\ & & \dots & \dots & \dots & \\ & & & 1 & -2 & 1 \end{pmatrix}. \quad (25)$$

3.3 Globální řešení datových a strukturálních rovnic pro obálku harmonického signálu

Stejně jako v případě základního filtru, také řešení těchto rovnic se sestaví pomocné kritérium z euklidovských norem budící funkce a zbytkového signálu ve tvaru

$$J = \mathbf{x}^H \mathbf{D} \mathbf{x} + (\mathbf{y}^T - \mathbf{x}^H \mathbf{C}^H)(\mathbf{y} - \mathbf{C} \mathbf{x}). \quad (26)$$

kde $\mathbf{D} = r^2 \mathbf{A}^T \mathbf{A}$ je pomocná čtvercová matice zjednodušující zápis vzorců. Minimum kritéria (26) lze vypočítat například derivací podle reálné a imaginární složky vektoru $\mathbf{x} = \mathbf{u} - j\mathbf{v}$. Výsledkem výpočtu je vektor minimalizující kritérium (26):

$$\mathbf{x} = \mathbf{B}^{-1} \mathbf{C}^H \mathbf{y}, \quad (27)$$

kde $\mathbf{B} = \mathbf{D} + \mathbf{E}$.

Násobením vektoru měřených dat \mathbf{y} zleva diagonální maticí \mathbf{C}^H se posouvá frekvence sledované harmonické složky signálu do nulové frekvence, tj. takto korigovaný signál se stane stejnosměrnou složkou signálu. Platí

$$\mathbf{C}^H \mathbf{y} = \mathbf{C}^H (\mathbf{C} \mathbf{x} + \boldsymbol{\eta}) = \mathbf{C}^H \mathbf{C} \mathbf{x} + \mathbf{C}^H \boldsymbol{\eta} = \mathbf{x} + \mathbf{C}^H \boldsymbol{\eta}. \quad (28)$$

Násobením obou diagonálních matic \mathbf{C}^H a \mathbf{C} se kompenzuje velikost průběžné fáze na nulu, a výsledkem násobení je jednotková matice. Vektor $\mathbf{C}^H \mathbf{y}$ je složen z vektoru komplexní obálky signálu \mathbf{x} a transformovaného zbytkového signálu. Filtr plní funkci dolnopropropustné filtrace časového průběhu obálky. Jeho frekvenční charakteristika závisí na matici \mathbf{B} . Účinek filtrace lze ovládat volbou velikosti váhového koeficientu.

K výpočtu lze použít dříve uvedených programů s drobnými úpravami, které spočívají v náhradě jedné diagonály matice \mathbf{A} , která je zaplněna hodnotami $-c(n)$, konstantou -2 a v korekci vektoru měřených dat diagonální maticí \mathbf{C}^H . Násobení \mathbf{C}^H a \mathbf{y} lze provést postupem násobení prvku z diagonály prvkem vektoru \mathbf{y} . Diagonálu zmíněné matice lze vypočítat výrazem $\exp(-j*2*pi*cumsum(f)*dt)$, přičemž funkce `cumsum` nahrazuje integraci úhlové rychlosti $2*pi*f$. Proměnná `dt` je vzorkovací interval. Vektor $\mathbf{C}^H \mathbf{y}$ se vypočítá příkazem

$$yy = \exp(-j*2*pi*cumsum(f)*dt) .* y;$$

Násobením prvku s prvkem zajišťuje operátor tečky předcházející symbol násobení (`.*`). Výsledek výpočtu podle vzorce (29) je komplexní vektor \mathbf{x} . Absolutní hodnota tohoto

vektoru je však poloviční oproti výchozímu vektoru, což je způsobeno komplexním tvarem harmonického signálu. Korigovaný vektor \underline{yy} je proto vynásoben 2. Časový průběh komplexní obálky signálu je funkce $\text{abs}(x)$. Výsledný program má číslo 2.

Program pro Vold-Kalmanův filtr č.2

```
function x=MyVoldKalman2(y,dt,f,r,filtord)
% Funkce x= MyVoldKalman2(y,dt,f, r, filtord) pro vypocet
% x=inv(r*r*A'*A+E)*C*y,
% kde C = diag(exp(-j*fi(1)),...,exp(-j*fi(N))), fi je faze,
% která je vypočtena podle vzorce 2*pi*cumsum(f)*dt
% Vstupy:
% y vektor vstupnich dat
% f vektor frekvence f v Hz (y a f maji stejny pocet prvku)
% dt vzorkovaci perioda v sec
% r vahovy koeficient
% filtord rad filtru
% Vystup:
% x vektor minimalizuje funkci J= x'A'*K'*K*A*x+(y'-x'C')*(y-Cx)

N = max(size(y));
if filtord==1,
    NR = N-2;
else
    NR = N-3;
end;
e = ones(NR,1);
if filtord==1,
    A = spdiags([e -2*e e],0:2, NR, N);
else
    A = spdiags([e -3*e 3*e -e],0:3, NR, N);
end;
AA = r*r*A'*A +speye(N);
yy = exp(-j*2*pi*cumsum(f)*dt).*y;
x = 2*AA\yy;
```

4 Vold-Kalmanův filtr pro extrahování obecného počtu složek signálů

4.1 Soustava rovnic k řešení

Pro obecný počet P extrahovaných složek má kritérium pro volbu velikostí všech neznámých vektorů obecný tvar

$$J = \sum_{k=1}^P \boldsymbol{\varepsilon}_{Kk}^T \boldsymbol{\varepsilon}_{Kk} + \boldsymbol{\eta}^T \boldsymbol{\eta} = \sum_{k=1}^P r^2 \mathbf{x}_k^H \mathbf{A}^T \mathbf{A} \mathbf{x}_k + \left(\mathbf{y}^T - \sum_{k=1}^P \mathbf{x}_k^H \mathbf{C}_k^H \right) \left(\mathbf{y} - \sum_{k=1}^P \mathbf{C}_k \mathbf{x}_k \right). \quad (29)$$

Podmínka minima kritéria vyplývá z nulové derivace kritéria (29) podle každého vektoru ze skupiny $\mathbf{x}_i^H, i=1, \dots, P$. Toto derivování se nazývá konjugované. Pravidla konjugovaného derivování podle vektorů lze nalézt např. v příloze monografie [Haykin 1996]. Po substituci $\mathbf{B}_k = r^2 \mathbf{A}^T \mathbf{A} + \mathbf{E}$ platí

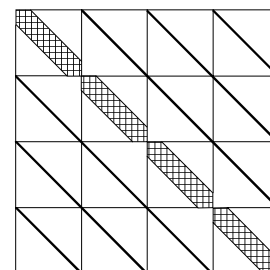
$$\frac{\partial J}{\partial \mathbf{x}_i^H} = \mathbf{B}_i \mathbf{x}_i + \mathbf{C}_i^H \sum_{\substack{k=1 \\ k \neq i}}^P \mathbf{C}_k \mathbf{x}_k - \mathbf{C}_i^H \mathbf{y} = \mathbf{0}, \quad i = 1, \dots, P. \quad (30)$$

Při úpravě posledního výrazu bylo využito vlastností prvků vektoru \mathbf{C}_i , které jsou rovny $\exp(j\theta_i(n))$. Je zřejmé, že pak platí $\mathbf{C}_i^H \mathbf{C}_i = \mathbf{E}$.

Matice soustavy rovnic (30) včetně vektoru neznámých veličin a vektoru pravé strany má následující strukturu.

$$\mathbf{B} = \begin{pmatrix} \mathbf{B}_1 & \mathbf{C}_1^H \mathbf{C}_2 & \dots & \mathbf{C}_1^H \mathbf{C}_P \\ \mathbf{C}_2^H \mathbf{C}_1 & \mathbf{B}_2 & \dots & \mathbf{C}_2^H \mathbf{C}_P \\ \dots & \dots & \dots & \dots \\ \mathbf{C}_P^H \mathbf{C}_1 & \mathbf{C}_P^H \mathbf{C}_2 & \dots & \mathbf{B}_P \end{pmatrix}, \quad \mathbf{x} = \begin{pmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \\ \dots \\ \mathbf{x}_P \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} \mathbf{C}_1^H \mathbf{y} \\ \mathbf{C}_2^H \mathbf{y} \\ \dots \\ \mathbf{C}_P^H \mathbf{y} \end{pmatrix}. \quad (31)$$

Pásové matice \mathbf{B}_i s reálnými prvky jsou symetrické a pozitivně definitní a protože pro komplexní diagonální matice $\mathbf{C}_i^H \mathbf{C}_j$ platí $\mathbf{C}_i^H \mathbf{C}_j = (\mathbf{C}_j^H \mathbf{C}_i)^H$, je matice \mathbf{B} soustavy rovnic (30) hermitovská. Struktura nenulových prvků matice pro $P = 4$ je znázorněna na obr. 2. Jednotlivé bloky jsou čtvercové matice o počtu řádků dosahujících desítek tisíc. Blokované matice na hlavní diagonále jsou pásové matice s několika nenulovými diagonálami (mezní počet několik desítek). Mimo hlavní diagonálu jsou blokované matice diagonální.



Obr. 2 Struktura matice soustavy rovnic (31)

4.2 Iterační řešení soustavy rovnic

Podle názoru expertů v oblasti lineární algebry (prof. Dostál z VŠB – TU Ostrava) je nepraktické hledat explicitní vzorce pro výpočet soustavy rovnic se strukturou naznačenou na obr. 2, ale je výhodnější řešit tuto soustavu rovnic iteračním postupem, který bude těžit z vlastností matice soustavy. Tento postup pro Vold-Kalmanův filtr dříve navrhli také Feldbauer, Ch. & Holdrich, R. (viz seznam literatury). Iteračních metod pro řešení soustav lineárních rovnic je velmi mnoho. Monografie popisující iterační řešení soustav lineárních rovnic (Saad: Iterative Methods for Sparse Linear Systems) nebo dokumentace k Matlabu označují jako zvláště vhodnou a navíc oblíbenou metodu pro řídké pozitivně definitní symetrické matice (symmetric positive definite – SPD) matice tzv. Preconditioned Conjugate Gradient (PCG) Algorithm.

Metoda PCG je kombinací metody CG (Conjugate Gradient) a techniky použití předpodmínkové matice (Preconditioner matrix). Metoda CG byla poprvé popsána autory Hestenes, M.R. & Stiefel E. v roce 1952.

V MATLABu je obsažena metoda PCG, která má různé varianty volání. Pro verzi, kdy je zadána předpodmínková matice \mathbf{M} nebo její rozklad $\mathbf{M} = \mathbf{M}_1 \mathbf{M}_2$ jsou varianta volání této funkce následující

$$\mathbf{x} = \text{pcg}(\mathbf{A}, \mathbf{b}, \text{TOL}, \text{MAXIT}, \mathbf{M}) \text{ nebo } \mathbf{x} = \text{pcg}(\mathbf{A}, \mathbf{b}, \text{TOL}, \text{MAXIT}, \mathbf{M}_1, \mathbf{M}_2)$$

resp. s počátečním odhadem \mathbf{x}_0

$$\mathbf{x} = \text{pcg}(\mathbf{A}, \mathbf{b}, \text{TOL}, \text{MAXIT}, \mathbf{M}, \mathbf{x}_0) \text{ nebo } \mathbf{x} = \text{pcg}(\mathbf{A}, \mathbf{b}, \text{TOL}, \text{MAXIT}, \mathbf{M}_1, \mathbf{M}_2, \mathbf{x}_0),$$

kde TOL je tolerance řešení a MAXIT je maximální počet iterací. Kromě výpočtu neznámého vektoru \mathbf{x} lze vypočítat také další parametry

$$[\mathbf{x}, \text{FLAG}, \text{RELRES}, \text{ITER}, \text{RESVEC}] = \text{pcg}(\mathbf{A}, \mathbf{b}, \text{TOL}, \text{MAXIT}, \mathbf{M}_1, \mathbf{M}_2, \mathbf{x}_0)$$

kde RELRES je relativní residuum $\text{NORM}(\mathbf{b} - \mathbf{A} * \mathbf{x}) / \text{NORM}(\mathbf{b})$, přičemž FLAG = 0, jestliže RELRES je menší nebo rovno než TOL, dále je vrácena hodnota skutečného počtu iterací ITER (ITER <= MAXIT) a vektor residuálních norem RESVEC včetně $\text{NORM}(\mathbf{b} - \mathbf{A} * \mathbf{x}_0)$.

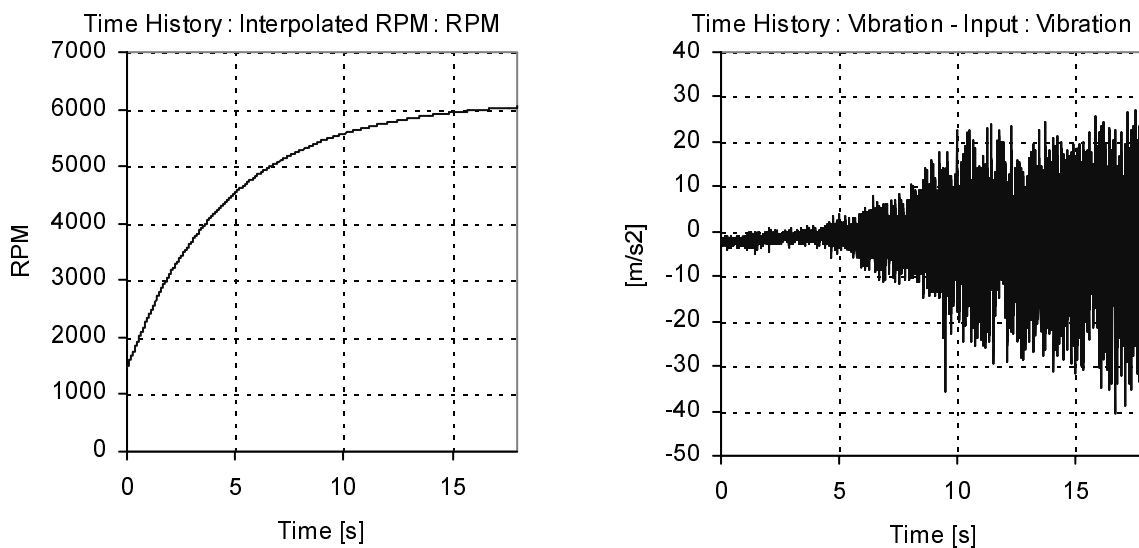
Za předpodmínkovou maticí soustavy lze zvolit

$$\mathbf{M} = \begin{pmatrix} \mathbf{B}_1 & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{B}_2 & \dots & \mathbf{0} \\ \dots & \dots & \dots & \dots \\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{B}_p \end{pmatrix}. \quad (32)$$

Vzhledem k tomu, že matice \mathbf{B} je pásová a pětidiagonální, bude řešení předpokládané soustavy $\mathbf{M}\mathbf{u} = \mathbf{b}$ snadné metodou Choleskiho rozkladu na horní a dolní trojúhelníkovou matici. Výsledek výpočtu \mathbf{u} předpokládané soustavy rovnic je použit při hledání aproximace řešení metodou CG.

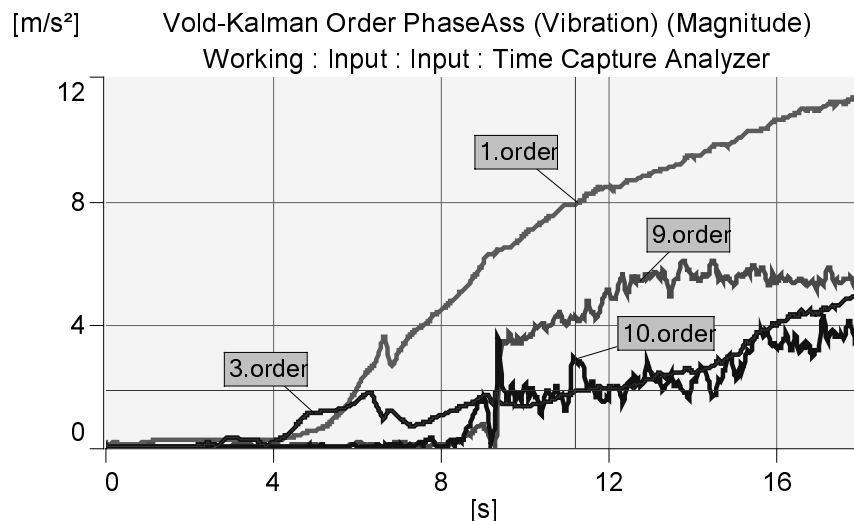
5 Příklad

Demonstrační projekt R-UMotor.pls, který je součástí software PULSE, obsahuje měření vibrací a otáček během rozběhu motoru s průběhem znázorněným na obr. 3.



Obr. 3 Časový průběh otáček a vibrací při rozběhu motoru

Cílem řádové analýzy bylo určit amplitudu harmonických složek, které jsou zvoleným násobkem frekvence otáček motoru. V analyzátoru PULSE šlo o volbu výstupu (Output) na Phase assigned order. Výsledek výpočtu amplitudy v analyzátoru PULSE je diagram na obr. 4.



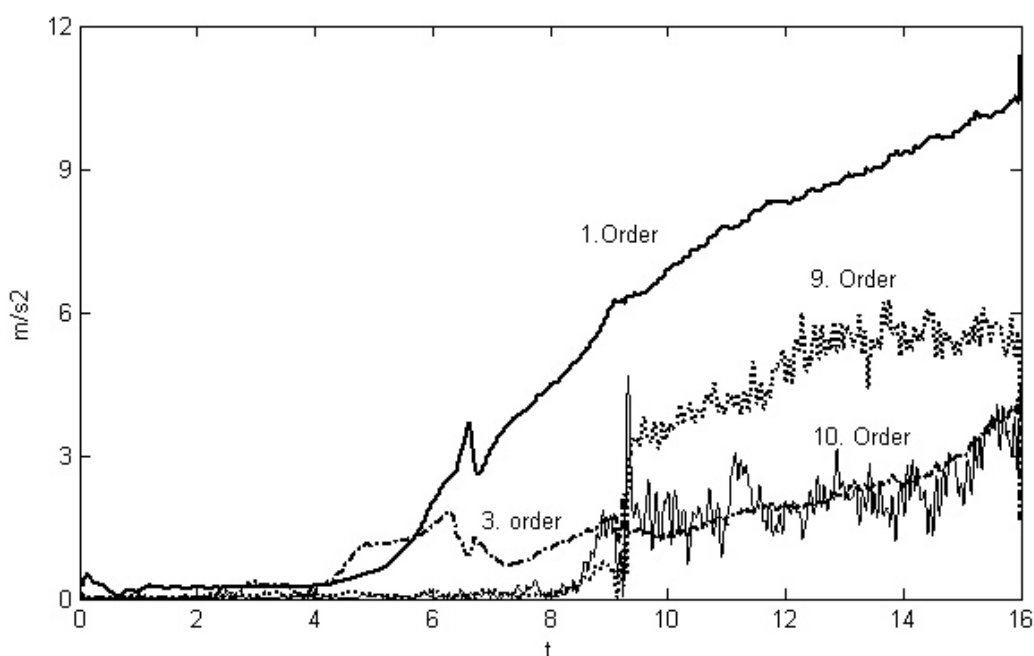
Obr. 4 Výsledek výpočtu Vold-Kalmanovy filtrace v analyzátoru PULSE

Popiska Order označuje efektivní hodnoty složek o frekvenci dané určenými násobky (1, 3, 9, 10) frekvence otáček. Jde tedy o velikosti hodnot obálky podělené odmocninou ze dvou protože analyzátor PULSE preferuje efektivní hodnoty (RMS) před obálkou, tj. amplitudou signálu.. Pro výpočet obálky je zvolen dvoupólový filtr (druhého řádu).

Při řádové filtraci musí být kolem sledované frekvence ještě určité propustné pásmo, aby byl zachycen fázový a amplitudový modulační signál, který je obsažen v postranních pásmech složky. V daném příkladě byla šířka propustného pásma kolem základní harmonické složky (jednonásobek otáček) zvolena 10%. Pro vyšší násobky pak úměrně menší šířka. Data byla 100krát decimována (vykresluje se jen každý stý vzorek).

Řízení šířky propustného pásma Vold-Kalmanova filtru nebylo v tomto krátkém příspěvku popisováno, i když je již v Matlabu vypracován příslušný program.

Na obr. 5 je výsledek výpočtu Vold-Kalmanovy filtrace v prostředí Matlab programem, který je označen č. 2 s příslušným rozšířením pro kontrolu šířky propustného pásma filtru. Podmínky výpočtu byly totožné, avšak data nebyla decimována. Podobnost obou grafů dokládá shodnost algoritmu výpočtu.



Obr. 4 Výsledek výpočtu Vold-Kalmanovy filtrace v Matlabu

6 Závěr

Vold-Kalmanova filtrace je nový nástroj pro řádovou filtraci signálů z točivých strojů, u kterých je součástí záznamu také signál otáček. Úloha filtrace vede na řešení rozsáhlých soustav rovnic obsahujících až statisíce neznámých. Matlab s možností paměťově šetrného ukládání řídkých (sparse) matic je velmi vhodným prostředkem pro zpracování zmíněných signálů nebo ke studiu vlastností popsaného algoritmu.

Literatura

Leuridan, J. & Van der Auweraer, H. & Vold, H. The analysis of non-stationary signals, LMS International, Application Notes

Leuridan, J. & Van der Auweraer, H. & Vold, H. The analysis of non-stationary signals, Sound and Vibration, August 1994

- Vold, H. & Leuridan, J. High Resolution Order Tracking at Extreme Slew Rates, Using Kalman Tracking Filter, SAE Paper 931288, 1993
- Blough, J. R. & Brown, D.L. & Vold, H. The Time Variant Discrete Fourier Transform as an Order Tracking Method, SAE Paper 972006, 1997
- Vold, H. & Mains, M. & Brown, D.L. Theoretical Foundations for High Performance Order Tracking with Vold-Kalman Tracking Filter, SAE Paper 972007, 1997
- Vold, H. & Herlufsen, H. & Mains, M. & Corwin-Renner, D. Multi Axle Order Tracking with Vold-Kalman Tracking Filter, Sound & Vibration 1997
- Herlufsen, H. & Gade, S. & Konstantin-Hansen, H. & Vold, H. Characteristics of the Vold/Kalman Order Tracking Filter, Sound & Vibration 1997
- Gade, S. & Herlufsen, H. Konstantin-Hansen, H. & Vold, H. Characteristics of the Vold-Kalman Order Tracking Filter. Brüel & Kjaer Technical Review, No.1 - 1999.
- Haykin S. Adaptive Filter Theory, Third Edition, Prentice-Hall International, INC., New Jersey 1996, ISBN 0-13-397985-7
- Fiedler M. Speciální matice a jejich použití v numerické matematice, TKI SNTL Praha 1981
- Feldbauer, Ch. & Holdrich, R. Realisation of a Vold-Kalman Tracking Filter – A Least Square Problem, Proceedings of the COST G-6 Conference on Digital Audio Effects (DAFX-000, Verona Italy, December 7-9, 2000
- Randall, R.B. Frequency Analysis, Bruel & Kjaer, Denmark, Revision September 1987, ISBN 87-87355-07-8
- Saad, Y. Iterative Methods for Sparse Linear Systems, Second Editions with Corrections, January 3rd 2000.
- Hestenes, M.R. & Stiefel E. Methods of Conjugate Gradients for Solving Linear Systems, Journal of Research of the National Bureau of Standards, Vol. 49, No. 6, December 1952. Research Paper 2379.

Poděkování Tento výzkum je prováděn na katedře Automatizační techniky a řízení VŠB Technické university v Ostravě jako část výzkumného záměru číslo CEZ J17/03:272300011, který je podporován Ministerstvem školství a mládeže České republiky.