

FREQUENCY SPECTRUM ESTIMATION BY AUTOREGRESSIVE MODELING

J. Tůma^{*}

Summary: *The paper deals with methods for frequency spectrum estimation by autoregressive modeling. Estimate of the autoregressive model parameters is the first step in this way of spectral analysis. The method used to do it are based on solving the system of normal equations, resulting from the least square method, or on solving the Yule-Walker equation and on Burg's method. The least square method uses the Gram-Schmidt orthogonalisation and its modification. The Yule-Walker equations are solved using Cholesky's factorization and Levinson's iterative method. The AR model parameters determine the transfer function of a linear system with white noise as an input signal. The frequency transfer function and the model error variance determine the frequency spectrum.*

1. Úvod

K výpočtu frekvenčních spekter signálů jsou obecně k dispozici (Orfanidis, 1988):

- neparametrické metody
- parametrické metody
- tzv. subspace methods.

Neparametrické metody frekvenční analýzy jsou založeny na pásmových filtrech pro zaznamenaný signál. Mezi tyto metody patří také Fourierova transformace. Tento způsob výpočtu spektra signálu je velmi rozšířen a mezi technikou je povědomí, že nic jiného k výpočtu spektra neexistuje.

Parametrické metody jsou naproti tomu založeny na výpočtu autoregresního modelu časové řady vzorků, které jsou považovány za výstup lineárního dynamického systému se vstupním signálem typu bílého šumu. Spektrum výstupního signálu je pak dáno frekvenční přenosovou funkcí této soustavy a faktorem jejího zesílení, který je dán rozptylem vstupního bílého šumu. K výpočtu autoregresního modelu lze použít několik metod, z nichž metoda nejmenších čtverců a řešení soustavy Youle-Walkerových rovnic jsou popsány blíže v tomto referátu spolu se dvěma způsoby řešení soustavy Youle-Walkerových rovnic pomocí Choleskyho rozkladu nebo Levinsonova rekurzivního algoritmu.

Anglicky pojmenované „subspace methods“ jsou charakteristické vysokým rozlišením také pro velmi malé počty vzorků, kdy výpočet spektra užitím Fourierovy transformace obsahuje velmi málo složek a tudíž spektrum je s malým rozlišením na frekvenční ose. Princip těchto

^{*} Prof. Ing. Jiří Tůma, CSc.: Fakulta strojní, VŠB – Technická univerzita Ostrava; 17. listopadu 15; 708 33 Ostrava; tel.: +420 59 699 3482, fax: +420 59 699 3482; e-mail: jiri.tuma@vsb.cz

metod spočívá na rozkladu korelační matice signálu na vlastní vektory. Příkladem takovéto metody je například MUSIC (Multiple Signal Classification). Popis tohoto postupu výpočtu není předmětem referátu.

2. Autoregresní model

Jak již bylo uvedeno, parametrické metody se opírají o výpočet parametrů autoregresního modelu signálu. Autoregresní model popisuje závislost aktuálního vzorku y_t signálu na předchozích hodnotách vzorků $y_{t-1}, y_{t-2}, \dots, y_{t-M}$, tj.

$$y_t = a_1 y_{t-1} + a_2 y_{t-2} + \dots + a_M y_{t-M} + e_t, \quad (1)$$

kde t představuje diskrétní čas (0, 1, 2, ...), M je řád modelu, $a_1, a_2, \dots, a_M$ ($a_M \neq 0$) jsou parametry a e_t je náhodná chyba typu nekorelovaného šumu

$$E\{e_t^2\} = \sigma^2 > 0, E\{e_t e_{t+k}\} = 0, k \neq 0. \quad (2)$$

Z tvaru diferenční rovnice je zřejmé, že aktuální velikost výstupní veličiny je součtem váženého průměru minulých hodnot o počtu, který je roven řádu soustavy, a náhodné složky s vlastnostmi bílého šumu. Diferenční rovnice má tvar regresní rovnice s jednou náhodnou chybou. Regresní rovnice popisuje závislost aktuální velikosti výstupní veličiny na jejích minulých hodnotách. Tento typ modelu se nazývá autoregresní, zkratkou AR (AutoRegressive) model.

Ve speciální odborné literatuře o náhodných procesech se k zápisu diferenčních rovnic používá operátor posunutí q^{-1} . Zpoždění o jeden časový krok (jednotku času, tj. z času t na $t-1$) lze tedy zapsat jako $y_{t-1} = q^{-1} y_t$, resp. zpoždění o k vzorků $y_{t-k} = q^{-k} y_t$. Symbolický zápis diferenční rovnice se pak změní použitím tohoto operátoru na tvar

$$(1 - a_1 q^{-1} - a_2 q^{-2} - \dots - a_M q^{-M}) y_t = e_t, \quad (3)$$

resp. $A(q^{-1}) y_t = e_t$. Formální podíl výstupního vzorku již zmíněné soustavy a vstupního vzorku chyby představuje operátor, který lze označit za operátorový přenos soustavy

$$G_{ey}(q^{-1}) = \frac{y_t}{e_t} = \frac{1}{(1 - a_1 q^{-1} - a_2 q^{-2} - \dots - a_M q^{-M})}, \quad (4)$$

resp.

$$G_{ey}(q^{-1}) = \frac{y_t}{e_t} = \frac{1}{A(q^{-1})}. \quad (5)$$

Odstraněním záporných mocnin operátoru posunutí lze získat přenosovou funkci ve tvaru

$$G_{ey}(q) = \frac{y_t}{e_t} = \frac{z^M}{(q^M - a_1 q^{M-1} - a_2 q^{M-2} - \dots - a_{M-1} q - a_M)} \quad (6)$$

Podmínkou stability přenosové funkce je umístění kořenů polynomu v proměnné q ve jmenovateli přenosové funkce uvnitř jednotkové kružnice.

Operátor posunutí odpovídá zpoždění o periodu vzorkování T . Frekvenční přenosová funkce dopravního zpoždění formálně souvisí s operátorem posunutí podle vzorce

$$q^{-1} = \exp(-j\omega T), \quad (7)$$

kde úhlová frekvence ω v rad/s a frekvence v Hz lze přepočítat podle vzorce $\omega = 2\pi f$. Substitucí proměnné q podle vzorce (7) v operátorové přenosové funkci (4) lineárního číslicového filtru lze dostat frekvenční přenosovou funkci, která odpovídá frekvenční přenosové funkci spojitého systému.

Na vstupu lineární soustavy s uvedeným přenosem je bílý šum s konstantní výkonovou spektrální hustotou ve frekvenčním pásmu od nuly do Nyquistovy frekvence $1/2T$. Celkový výkon signálu bílého šumu je σ^2 , tj. shodný s rozptylem chyby modelu (1). S ohledem na postup výpočtu tohoto celkového výkonu lze výkonovou spektrální hustotu bílého šumu určit podle vztahu $2\sigma^2 T$. Výkonová spektrální hustota na výstupu soustavy je dána součinem výkonové spektrální hustoty na vstupu soustavy a druhé mocniny absolutní hodnoty frekvenční přenosové funkce. Frekvenční přenos tedy je shodný až na konstantní faktor se spektrem signálu na výstupu soustavy. Těto vlastnosti je využito v parametrických metodách výpočtu frekvenčního spektra signálů. Výkonová spektrální hustota signálu je dána vzorcem

$$P(\omega) = \frac{2\sigma^2 T}{\left| 1 - \sum_{m=1}^M a_m \exp(-jm\omega T) \right|^2}. \quad (8)$$

Prvým krokem výpočtu je tedy určení parametrů autoregresního modelu. V tomto textu budou popsány jen některé metody výpočtu, avšak v demonstračním příkladu budou použity metody tři.. První představuje metodu nejmenších čtverců, druhá o Yule-Walkovy rovnice a třetí Burgovu metodu.

3. Metoda nejmenších čtverců

Cílem výpočtu je určit parametry modelu tak, aby rozptyl chyby modelu nabýval minima. Neznámé parametry jsou řešením přeuročené (více rovnic než neznámých) soustavy rovnic

$$\mathbf{A} \mathbf{a} = \mathbf{b}, \quad (9)$$

kde $\mathbf{A}$ je obecně obdélníková matice soustavy, jejíž sloupce obsahují vzorky signálu postupně posunuté o jeden vzorek

$$\mathbf{A} = \begin{pmatrix} y_0 & 0 & \dots & 0 \\ y_1 & y_0 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ y_{N-2} & y_{N-3} & \dots & y_{N-M-1} \end{pmatrix} \quad (10)$$

$\mathbf{a}$ je vektor neznámých parametrů a $\mathbf{b}$ je vektor pravých stran. Pro transponované vektory platí

$$\mathbf{a}^T = (a_1 \ a_2 \ \dots \ a_M), \quad \mathbf{b}^T = (y_1 \ y_2 \ \dots \ y_{N-1}). \quad (11)$$

Řešení této je dáno soustavou normálních rovnic, jejíž řešení je následující

$$\mathbf{a} = (\mathbf{A}^T \cdot \mathbf{A})^{-1} \mathbf{A}^T \mathbf{b} \quad (12)$$

a ve které se invertuje čtvercová matice místo obdélníkové matice $\mathbf{A}$. Pro metodu nejmenších čtverců se s výhodou používá QR rozklad matice

$$\mathbf{A} = \mathbf{Q}\mathbf{R} \quad (13)$$

kde $\mathbf{Q}$ je ortogonální matice (součin s transponovanou maticí má za výsledek jednotkovou matici) a $\mathbf{R}$ je horní trojúhelníková matice. Výpočet neznámého vektoru $\mathbf{a}$ podle (12) se pak zjednoduší na vzorec

$$\mathbf{a} = \mathbf{R}^{-1} \mathbf{Q}^T \mathbf{b}. \quad (14)$$

Ke QR rozkladu obecně obdélníkové matice lze použít Gram-Schmidtův algoritmus nebo jeho modifikovanou verzi a nebo Householderovou matici (Saad, 2000).

4. Řešení Yule-Walkerových rovnic

Autokorelační funkce obecného AR modelu se získá ze střední hodnoty levé a pravé strany rovnice modelu vynásobené vzorkem, který je posunut ve směru přírůstků času o τ vzorků

$$y_t y_{t+\tau} - a_1 y_t y_{t+\tau-1} - a_2 y_t y_{t+\tau-2} - \dots - a_M y_t y_{t+\tau-M} = y_t e_{t+\tau}. \quad (15)$$

Na pravé straně rovnice je součin náhodné chyby a vzorku signálu, který tato chyba nemůže ovlivnit. Střední hodnota tohoto součinu je z důvodu nezávislosti chyby v čase $t + \tau$ a chyb, které určily vzorek signálu v čase t , je nulová. Na levé straně rovnice jsou součiny vzájemně posunutých vzorků signálu. Střední hodnota

$$E\{y_t y_{t+\tau}\} - a_1 E\{y_t y_{t+\tau-1}\} - a_2 E\{y_t y_{t+\tau-2}\} - \dots - a_M E\{y_t y_{t+\tau-M}\} = E\{y_t e_{t+\tau}\} \quad (16)$$

představuje následující podmínku, kterou splňuje M po sobě jdoucích hodnot autokorelační funkce

$$R_{yy}(\tau) - a_1 R_{yy}(\tau-1) - a_2 R_{yy}(\tau-2) - \dots - a_M R_{yy}(\tau-M) = 0. \quad (17)$$

Vzhledem k tomu, že hodnoty autokorelační funkce pro vzájemně opačná posunutí jsou shodné, výsledkem rozpisu této rovnice pro velikosti posunutí od 1 do M je soustava Yule-Walkerových rovnic, jejichž řešením jsou neznámé parametry autoregresního modelu

$$\begin{aligned} R_{yy}(1) - a_1 R_{yy}(0) - a_2 R_{yy}(1) - \dots - a_M R_{yy}(M-1) &= 0 \\ R_{yy}(2) - a_1 R_{yy}(1) - a_2 R_{yy}(0) - \dots - a_M R_{yy}(M-2) &= 0 \\ \dots\dots\dots \\ R_{yy}(M-1) - a_1 R_{yy}(M-2) - a_2 R_{yy}(M-3) - \dots - a_M R_{yy}(1) &= 0 \\ R_{yy}(M) - a_1 R_{yy}(M-1) - a_2 R_{yy}(M-2) - \dots - a_M R_{yy}(0) &= 0. \end{aligned} \quad (18)$$

Tuto soustavu rovnic lze zapsat ve tvaru

$$\mathbf{A} \mathbf{a} = \mathbf{b}, \quad (19)$$

kde $\mathbf{A}$ je matice soustavy, která je složena z hodnot korelační funkce, a $\mathbf{b}$ je vektor pravých stran

$$\mathbf{A} = \begin{pmatrix} R_{yy}(0) & R_{yy}(1) & \dots & R_{yy}(M-1) \\ R_{yy}(1) & R_{yy}(0) & \dots & R_{yy}(M-2) \\ \dots & \dots & \dots & \dots \\ R_{yy}(M-1) & R_{yy}(M-2) & \dots & R_{yy}(0) \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} R_{yy}(1) \\ R_{yy}(2) \\ \dots \\ R_{yy}(M) \end{pmatrix}. \quad (20)$$

Matice soustavy je význačná tím, že prvky na jejich diagonálách jsou shodné

$$A_{i,j} = R_{yy}(j-i). \quad (21)$$

Tento typ matic se označuje jako Toeplitzův. Matice soustavy rovnic je symetrická, a proto k řešení této soustavy lze s výhodou použít Choleskyho rozkladu na trojúhelníkové matice, které lze snadno invertovat

$$\mathbf{A} = \mathbf{L} \mathbf{U} \quad (\mathbf{U} = \mathbf{L}^T), \quad (22)$$

kde $\mathbf{L}$ je dolní trojúhelníková $\mathbf{U}$ je horní trojúhelníková matice. Po substituci $\mathbf{U} \mathbf{a} = \mathbf{c}$ je vektor neznámých parametrů $\mathbf{a}$ dán vztahem

$$\mathbf{c} = \mathbf{L}^{-1} \mathbf{b} \Rightarrow \mathbf{a} = \mathbf{U}^{-1} \mathbf{c}. \quad (23)$$

Řešení soustavy je pomocí Choleskyho rozkladu vyžaduje větší množství operací než výhodnější Levinsonova rekurzivní metoda (Orfanidis, 1988). Pro popis algoritmu rekurzivního výpočtu je zavedena funkce představující levé strany tzv. Yule-Walkerových rovnic (17) ve tvaru, ve kterém jsou pravé strany nulové

$$g_p(k) = R_{yy}(k) - a_1 R_{yy}(k-1) - a_2 R_{yy}(k-2) - \dots - a_p R_{yy}(k-p). \quad (24)$$

Rozptyl chyby modelu řádu $p+1$ je určen rekurzivním vztahem

$$E_{p+1} = g_{p+1}(0) = g_p(0) - \gamma_{p+1} g_p(p+1) = (1 - \gamma_{p+1}^2) g_p(0) = (1 - \gamma_{p+1}^2) E_p, \quad (25)$$

kde

$$\gamma_{p+1} = \frac{g_p(p+1)}{g_p(0)}, \quad (26)$$

Počáteční volba odhadu koeficientu a rozptylu E_p pro $p=0$ je následující

$$a_{01} = 1, \quad E_0 = R_{yy}(0) = E(e_t^2). \quad (27)$$

Vzorec (25) potvrzuje, že střední hodnota čtverce chyby modelu s rostoucím řádem postupně klesá. Rekurzivní výpočet parametrů modelu je dán vztahy

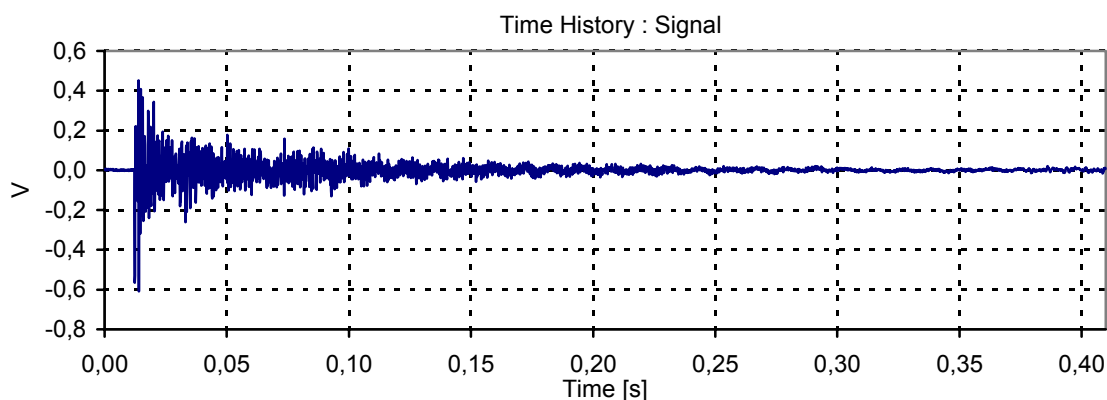
$$\begin{aligned} a_{p+1,m} &= a_{pm} - \gamma_{p+1} a_{p,p+1-m}, \quad 1 \leq m \leq p \\ a_{p+1,p+1} &= -\gamma_{p+1}. \end{aligned} \quad (28)$$

5. Příklad výpočtu spektra signálu

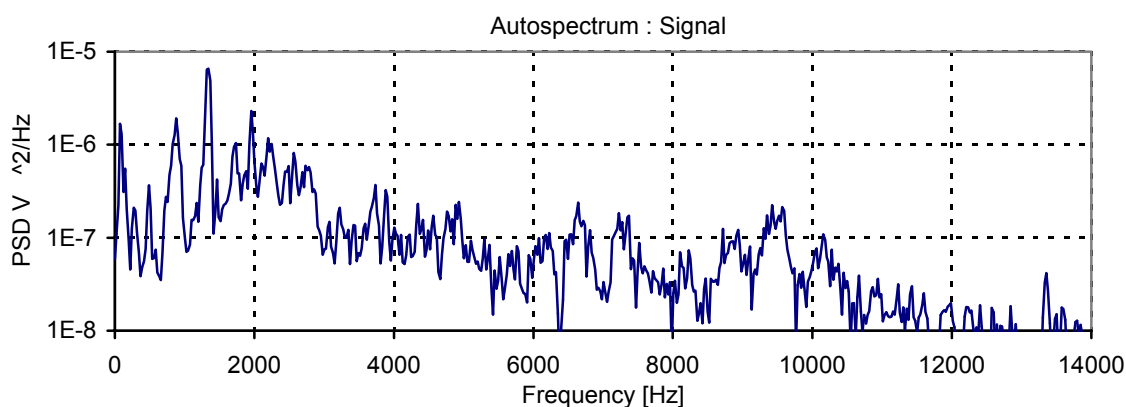
Pro demonstraci výpočtu spekter je použit testovací signál z měření impulsní odezvy, která obsahuje výrazné harmonické složky s aditivním náhodným šumem. Vzorkovací frekvence je 50 kHz a doba záznamu asi 0,4 s. Časový průběh tohoto signálu je znázorněn na obrázku 1.

Výkonová spektrální hustota je znázorněna na obrázku 2. Spektrum bylo vypočteno pro 800 složek průměrovaného spektra s překrytím záznamů 2/3.

Autoregresní model řádu 100 (AR 100) a řádu 400 (AR 400) byl vypočten metodou nejmenších čtverců (Least Square) ve variantě s Gram-Schmidtovou metodou (LS:GS) a s modifikovanou Gram-Schmidtovou metodou (LS:MGS) ortogonalizace a pomocí Yule-Walkerových rovnic ve variantě s Choleskyho rozkladem (YW:ChF) a s Levinsonovým rekurzivním algoritmem (YW:Lev). Poslední testovanou metodou je Burgův algoritmus (Burg's Method). Výsledky výpočtu jsou znázorněny v obrázcích 3 a 4.



Obrázek 1 Testovací signál 20500 vzorků, 50 kHz vzorkovací frekvence

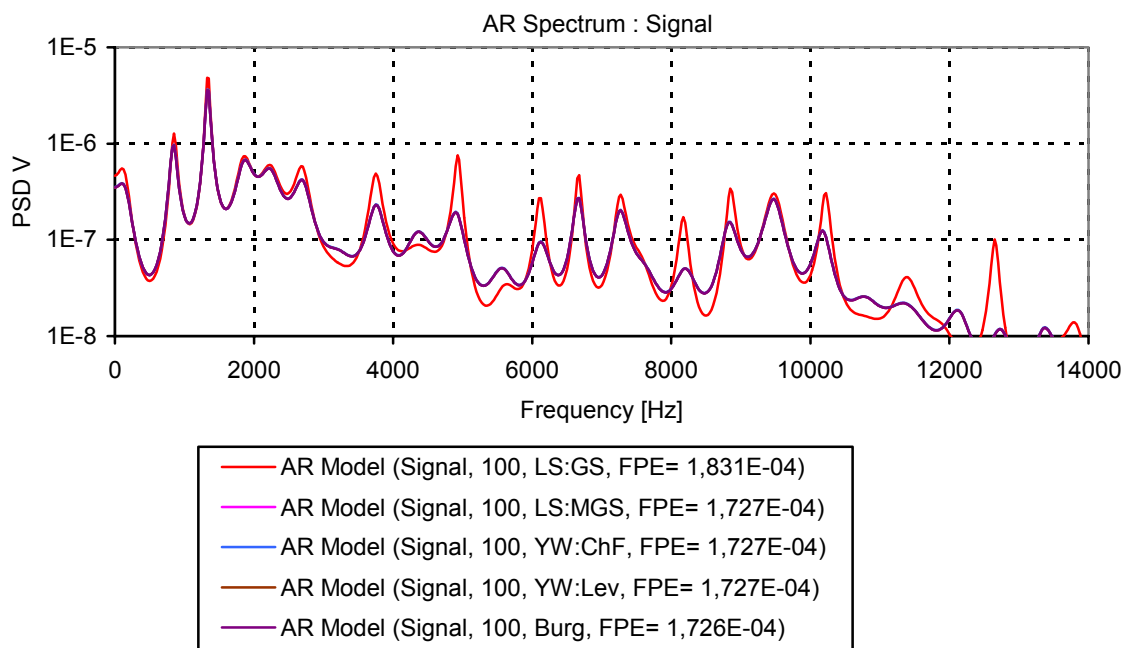


Obrázek 2 Výkonová spektrální hustota signálu, která byla vypočtena s užitím FFT

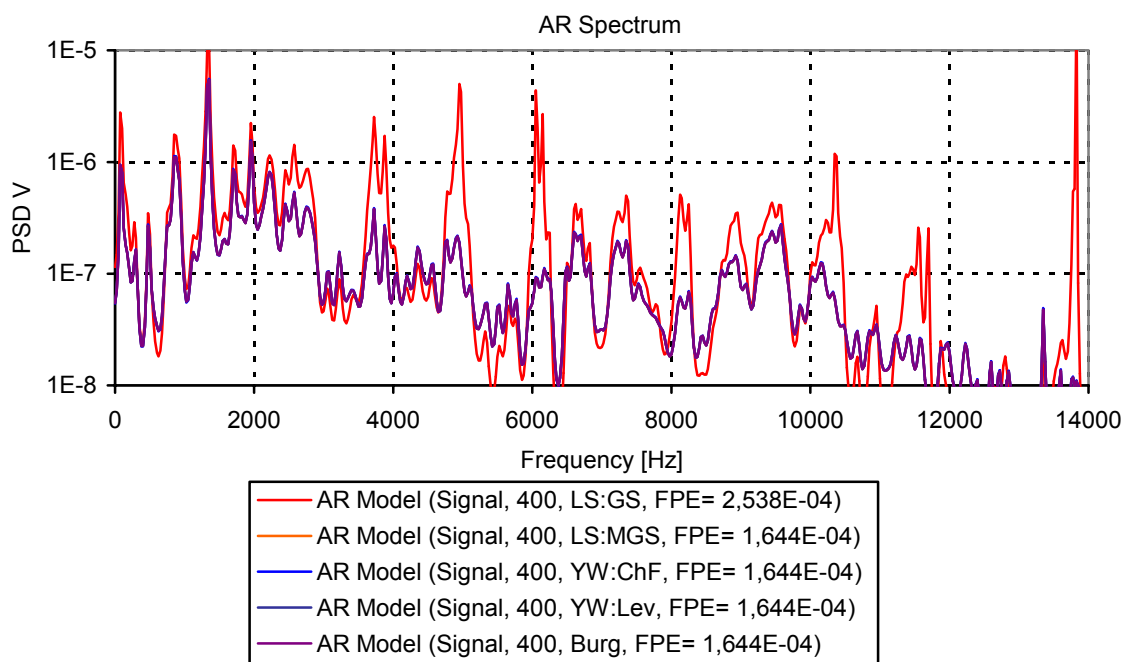
Součástí legendy grafu na obrázku 3 a dalších obrázků je údaj o chybě modelu. Jedná se o kritérium FPE (Final Prediction Error), které navrhl Akaike a které zohledňuje řád modelu M a počet vzorků N . Skutečný rozptyl chyby V modelu je korigován podle následujícího vzorce

$$FPE = \frac{1 + M/N}{1 - M/N} V \quad (29)$$

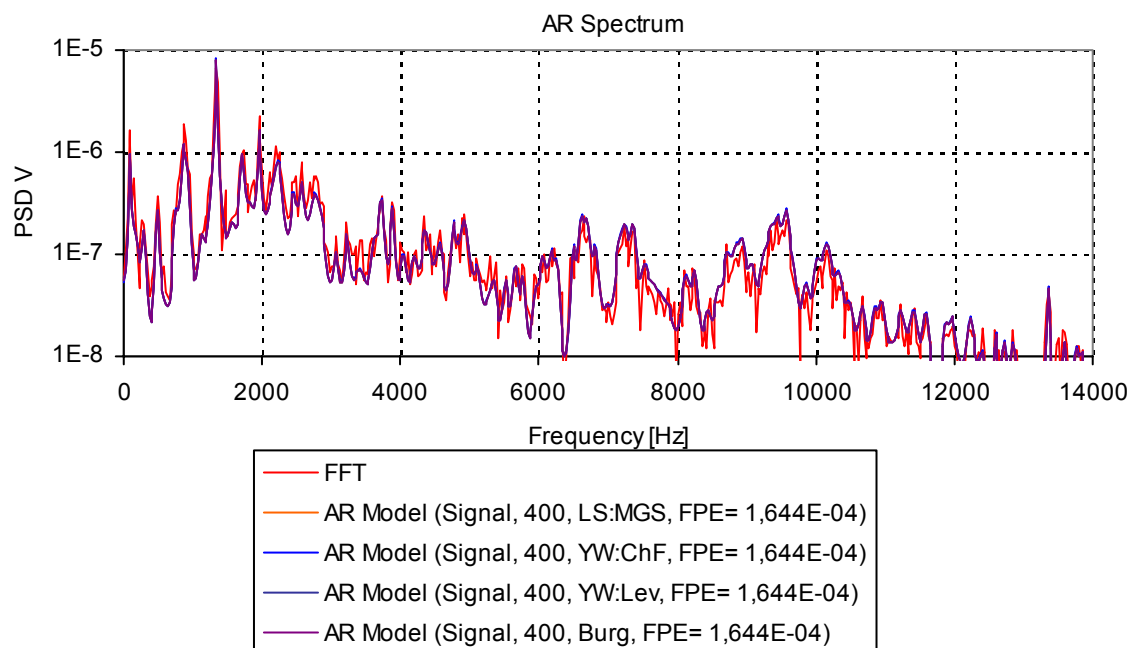
Kritérium FPE ukazuje, že až na metodu výpočtu podle nejmenších čtverců s Gram-Schmidtovou ortogonalizační metodou jsou ostatní metody rovnocenné. Obrázek 5 ukazuje téměř shodu s postupem výpočtu s využitím FFT. Problémem je volba řádu modelu. Je třeba uvést, že je to spíše otázka umění než přesného analytického řešení. Z tohoto hlediska je důležité, aby pro volbu řádu nebylo podstatného omezení.



Obrázek 3 Výkonové spektrální hustoty, které byly vypočteny s užitím AR 100



Obrázek 4 Výkonové spektrální hustoty, které byly vypočteny s užitím AR 400



Obrázek 5 Porovnání výsledků výpočtu výkonové spektrální hustoty s užitím FFT a některých metod pro AR 400

Doba výpočtu parametrů AR modelu z 20000 vzorků záznamu signálu z obrázku 1 je uvedena v tabulce 1. K výpočtům byl použit notebook HP Compaq nc6120, s procesorem 1,86 GHz a s pamětí 512 MB. Nejkratší výpočet je dosažen s použitím Yule-Walkerových rovnic a rekurzivním algoritmem, který navrhl Levinson (Orfanidis, 1988).

Tabulka 1

Metoda	Doba výpočtu	
	Řád AR modelu 100	Řád AR modelu 400
Least Square, Gram-Schmidt	18,91 s	288 s
Least Square, Modified Gram-Schmidt	24,95 s	387 s
Yule-Walker, Cholesky Factorization	0,44 s	2,92 s
Yule-Walker, Levinson	0,17 s	0,61 s
Burg's Method	2,05 s	7,72 s

6. Závěr

Výhoda AR modelu při výpočtu frekvenčních spekter spočívá v tom, že spektrum lze stanovit pro libovolný počet vzorků a na rozdíl od FFT výpočet není limitován na záznam o určité délce. Z testovaných metod výpočtu parametrů modelu byla nejrychlejší metoda na principu řešení Yule-Walkerových rovnic rekurzivním algoritmem, který navrhl Levinson.

Souhrnně lze uvést, že parametrická metoda spektrální analýzy je vhodná nejen pro náhodné signály, které popisují modely nízkého řádu, ale také pro signály obsahující

harmonické složky s aditivním šumem. Kromě uvedeného příkladu lze očekávat, že parametrická metoda najde hlavně uplatnění pro záznamy s malým počtem vzorků (stovky).

7. Poděkování

Výzkumné práce v oboru zpracování měření hluku a vibrací jsou podporovány Grantovou agenturou České republiky jako projekt č. 101/04/1530.

8. Literatura

Orfanidis, S.J. (1988) *Optimum signal processing: An Introduction*. Macmillan Publishing Company, London.

Saad, Y. (2000) *Iterative Methods for Sparse Linear Systems*. Second edition with corrections.

Press, W.H., Teukolsky, S.A., Vetterling, W.T. & Flannery B.P. *Numerical Recipes in FORTRAN 77: The art of scientific computing* (ISBN 0-521-43064-X), Cambridge University Press, 1986-1992.