

Lucrarea 9

Tehnici directe de proiectare a filtrelor digitale IIR

1. Noțiuni teoretice

Ca o alternativă la tehnica transformării unui filtru analogic în unul digital există metoda proiectării filtrelor IIR direct în domeniul discret, fără a face referire la cel analogic.

În primele două metode de proiectare ce vor fi prezentate, Padé și metoda celor mai mici pătrate, specificațiile și proiectarea sunt în domeniul timp. Există însă și tehnica de proiectare a filtrelor IIR în care optimizarea parametrilor filtrului se face în domeniul frecvență.

1.1. Metoda de aproximare Padé

În această metodă se presupune că răspunsul la impuls dorit $h_d[n]$ este specificat pentru $n \geq 0$. Filtrul pe care îl vom proiecta are funcția de sistem:

$$H(z) = \frac{\sum_{k=0}^M b_k z^{-k}}{1 + \sum_{k=1}^N a_k z^{-k}} = \sum_{n=0}^{\infty} h[n] z^{-n} \quad (1)$$

unde $h[n]$ este răspunsul la impuls. Filtrul are $L = M + N + 1$ parametri, și anume, coeficienții $\{a_k\}$ și $\{b_k\}$ ce pot fi aleși astfel încât să minimizeze un criteriu de eroare.

Criteriul celor mai mici pătrate este adesea folosit în problemele de optimizare. Să presupunem că se minimizează suma pătratelor erorilor:

$$\varepsilon = \sum_{n=0}^U [h_d[n] - h[n]]^2 \quad (2)$$

în raport cu parametrii $\{a_k\}$ și $\{b_k\}$, unde U este o limită superioară prestabilită în sumare.

În general, $h[n]$ este o funcție neliniară de parametrii filtrului și minimizarea lui ε implică rezolvarea unui set de ecuații neliniare. Dacă limita superioară a lungimii lui $h[n]$ se selectează ca fiind $U = L - 1$, este posibil a adapta perfect răspunsul dorit $h_d[n]$ pentru $0 \leq n \leq M + N$. Aceasta se poate realiza în felul următor: filtrul dorit este descris de ecuația cu diferențe

$$y[n] = -a_1 y[n-1] - a_2 y[n-2] - \dots - a_N y[n-N] + b_0 x[n] + b_1 x[n-1] + \dots + b_M x[n-M] \quad (3)$$

Dacă intrarea în filtru este impulsul unitate, $x[n] = \delta[n]$, rezultă că $y[n] = h[n]$ și:

$$h[n] = -a_1 h[n-1] - a_2 h[n-2] - \dots - a_N h[n-N] + b_0 \delta[n] + b_1 \delta[n-1] + \dots + b_M \delta[n-M] \quad (4)$$

Deoarece $\delta[n-k] = 0$ exceptând $n = k$, relația (4) devine:

$$h[n] = -a_1 h[n-1] - a_2 h[n-2] - \dots - a_N h[n-N] + b_n, \quad 0 \leq n \leq M \quad (5)$$

Pentru $n > M$, relația (4) devine:

$$h[n] = -a_1 h[n-1] - a_2 h[n-2] - \dots - a_N h[n-N], \quad n > M \quad (6)$$

Ecuatiile (5) și (6) pot fi folosite în determinarea coeficienților filtrului $\{a_k\}$ și $\{b_k\}$. Se impune $h[n] = h_d[n]$ pentru $0 \leq n \leq M + N$ și se folosește sistemul de ecuații (6) pentru a determina $\{a_k\}$. Apoi, aceștia se introduc în (5) pentru a determina coeficienții $\{b_k\}$.

Ecuatiile matriceale pentru determinarea coeficienților $\{a_k\}$ rezultă din relația (6):

$$\begin{bmatrix} h[M+1] \\ \vdots \\ h[M+N] \end{bmatrix} = \begin{bmatrix} h[M] & \dots & h[M-N+1] \\ \vdots & \dots & \vdots \\ h[M+N-1] & \dots & h[M] \end{bmatrix} \begin{bmatrix} a_1 \\ \vdots \\ a_N \end{bmatrix}$$

iar pentru determinarea coeficienților $\{b_k\}$ din relația (5):

$$\begin{bmatrix} b_0 \\ b_1 \\ \vdots \\ b_M \end{bmatrix} = \begin{bmatrix} 0 & 0 & \dots & 0 \\ h[0] & 0 & \dots & 0 \\ \vdots & \vdots & \dots & \vdots \\ h[M-1] & h[M-2] & \dots & h[M-N] \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_N \end{bmatrix} + \begin{bmatrix} h[0] \\ h[1] \\ \vdots \\ h[M] \end{bmatrix}$$

Astfel se obține o potrivire perfectă între $h[n]$ și $h_d[n]$ pentru primele L valori ale răspunsului la impuls. Această tehnică se numește de obicei **aproximarea Padé**.

Măsura în care această metodă permite obținerea de filtre acceptabile depinde, în parte, de numărul de coeficienți selectați. Evident, cu cât acesta este mai mare, cu atât va fi mai bună aproximarea. Acesta este un dezavantaj important al metodei și anume filtrul rezultat va avea mulți poli și multe zerouri, motiv pentru care folosirea sa în practică este limitată.

1.2. Proiectarea filtrelor digitale IIR prin metoda celor mai mici pătrate

În această metodă se adoptă un model pentru sistem și se determină parametrii modelului care minimizează în sensul celor mai mici pătrate eroarea dintre răspunsul sistemului real și răspunsul modelului.

Se presupune din nou că $h_d[n]$ este specificat pentru $n \geq 0$. Se începe cu cazul simplu, în care filtrul numeric ce urmează a fi proiectat conține numai poli, adică:

$$H(z) = \frac{b_0}{1 + \sum_{k=1}^N a_k z^{-k}} \quad (7)$$

Fie conectarea în cascadă a filtrului dorit $H_d(z)$ cu inversul $\frac{1}{H(z)}$, care este un filtru numai cu zerouri, ca în figura 1:

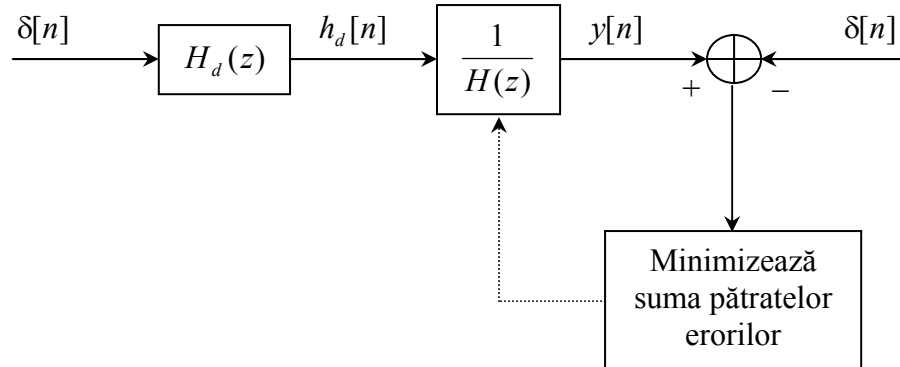


Figura 1 Proiectarea filtrului invers prin metoda celor mai mici pătrate

Se presupune că la intrarea cascadei din figura 1 se aplică excitația $\delta[n]$ astfel încât intrarea în sistemul invers este $h_d[n]$ și ieșirea $y[n]$. Ideal, $y_d[n] = \delta[n]$. Ieșirea reală $y[n]$ este:

$$Y(z) = H_d(z) \cdot \frac{1}{H(z)} = \frac{1}{b_0} H_d(z) \left(1 + \sum_{k=1}^N a_k z^{-k} \right) \quad (8)$$

$$y[n] = \frac{1}{b_0} \left(h_d[n] + \sum_{k=1}^N a_k h_d[n-k] \right)$$

Condiția $y_d[0] = y[0] = 1$ este satisfăcută prin alegerea $b_0 = h_d[0]$. Pentru $n > 0$, $y[n]$ reprezintă eroarea dintre ieșirea dorită $y_d[n] = 0$ și ieșirea reală. Parametrii $\{a_k\}$ vor fi selectați astfel încât să minimizeze suma pătratelor secvenței de eroare:

$$\varepsilon = \sum_{n=1}^{\infty} y^2[n] = \frac{\sum_{n=1}^{\infty} \left(h_d[n] + \sum_{k=1}^N a_k h_d[n-k] \right)^2}{h_d^2[0]} \quad (9)$$

Minimul acestei mărimi se obține pentru coeficienții $\{a_k\}$ rezultați din egalarea cu zero a derivatei lui ε în raport cu $\{a_k\}$, de unde se obține un set de ecuații liniare de forma:

$$\varepsilon' = \frac{\left(\sum_{n=1}^{\infty} \left(h_d^2[n] + 2h_d[n] \sum_{k=1}^N a_k h_d[n-k] + \sum_{k=1}^N a_k h_d[n-k] \sum_{l=1}^N a_l h_d[n-l] \right) \right)'}{h_d^2[0]} = 0 \Rightarrow$$

$$\Rightarrow \sum_{n=1}^{\infty} \left(2h_d[n] h_d[n-k] + \sum_{l=1}^N a_l h_d[n-k] h_d[n-l] + \sum_{k=1}^N a_k h_d[n-k] h_d[n-l] \right) = 0 \Leftrightarrow$$

$$\Leftrightarrow r_{dd}[k] + \sum_{l=1}^N a_l r_{dd}[k-l] = 0, \quad k = 1, 2, \dots, N$$

$$\sum_{l=1}^N a_l r_{dd}[k-l] = -r_{dd}[k], \quad k = 1, 2, \dots, N \quad (10)$$

unde $r_{dd}[k]$ este secvența de autocorelație a lui $\{h_d[n]\}$, definită ca:

$$r_{dd}[k] = \sum_{n=0}^{\infty} h_d[n] h_d[n+k] \quad (11)$$

Setul de ecuații (10) poate fi scris matricial:

$$[R_{dd}][a] = [r_{dd}] \quad (12)$$

unde $[R_{dd}]$ este matricea de corelație, cu elementele $\{r_{dd}[n]\}$, de dimensiune $N \times N$, $[a]$ este vectorul de dimensiune $N \times 1$ al coeficienților filtrului și $[r_{dd}]$ este un vector de dimensiune $N \times 1$ cu elementele $\{-r_{dd}[n]\}$.

Setul de ecuații (12) se rezolvă ușor cu ajutorul unui calculator ce efectuează calcul matricial, soluția fiind exprimată ca:

$$[a] = [R_{dd}]^{-1}[r_{dd}] \quad (13)$$

Metoda descrisă se numește *metoda de proiectare a filtrului invers pe baza celor mai mici pătrate*.

Pentru o problemă de proiectare particulară, răspunsul la impuls dorit $h_d[n]$ este specificat pentru un set finit de puncte; fie acesta $0 \leq n \leq L$, cu $L \gg N$. Într-un astfel de caz, secvența de corelație $r_{dd}[k]$ se poate calcula din secvența finită $h_d[n]$:

$$\hat{r}_{dd}[k-l] = \sum_{n=0}^{L-|k-l|} h_d[n] h_d[n+k-l], \quad 0 \leq k-l \leq N \quad (14)$$

aceste valori putând fi folosite pentru a forma pe $[R_{dd}]$ și $[r_{dd}]$. Astfel, setul de ecuații liniare (12) devine:

$$[\hat{R}][a] = [\hat{r}] \quad (15)$$

$$[a] = [\hat{R}]^{-1}[\hat{r}] \quad (16)$$

O metodă alternativă de rezolvare a problemei de aproximare numai cu poli se bazează pe conceptul de *predicție liniară pe baza celor mai mici pătrate*. După cum se arată în figura 2, ieșirea filtrului numai cu poli la un impuls $\delta[n]$ este:

$$y[n] = -\sum_{k=1}^N a_k h[n-k] + b_0 \delta[n] \quad (17)$$

sau, echivalent,

$$h[n] = -\sum_{k=1}^N a_k h[n-k] + b_0 \delta[n], \quad n = 0, 1, \dots \quad (17')$$

Răspunsul dorit este $h_d[n]$. Deoarece $h[0] = b_0$, se impune $b_0 = h_d[0]$. Pentru $n \geq 1$, relația (17) devine:

$$h[n] = -\sum_{k=1}^N a_k h[n-k] \quad (18)$$

condiția inițială fiind $h[n] = 0$ pentru $n < 0$.

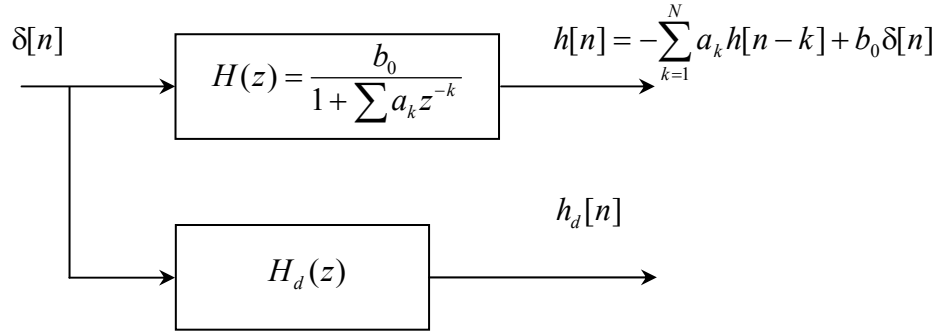


Figura 2: Proiectarea filtrului prin metoda celor mai mici pătrate, pe baza predicției liniare

Dacă $H_d(z)$ este un filtru numai cu poli, atunci:

$$h_d[n] = -\sum_{k=1}^N a_k h_d[n-k], \quad n \geq 1 \quad (19)$$

dar aceasta nu se întâmplă de obicei. Combinația liniară din membrul drept al relației (19) poate fi considerată ca un estimat al lui $h_d[n]$, adică:

$$\hat{h}_d[n] = -\sum_{k=1}^N a_k h_d[n-k], \quad n \geq 1 \quad (20)$$

$\hat{h}_d[n]$ se numește valoarea predicției liniare a lui $h_d[n]$. Suma pătratelor erorii de predicție dintre $h_d[n]$ și $\hat{h}_d[n]$ este:

$$\varepsilon = \sum_{n=1}^{\infty} (h_d[n] - \hat{h}_d[n])^2 = \sum_{n=1}^{\infty} \left(h_d[n] + \sum_{k=1}^N a_k h_d[n-k] \right)^2 \quad (21)$$

Expresia din relația (21) este chiar aceeași funcție eroare rezultată din setul de ecuații (10), ceea ce înseamnă că *metoda de predicție liniară pe baza celor mai mici pătrate* conduce la același rezultat ca metoda de proiectare a filtrului invers pe baza celor mai mici pătrate.

Metoda predicției pe baza celor mai mici pătrate poate fi extinsă la o aproximare a lui $H_d(z)$ care conține poli și zerouri. Dacă filtrul $H(z)$ care aproximează $H_d(z)$ are atât poli cât și zerouri, atunci răspunsul său la un impuls $\delta[n]$ este:

$$h[n] = -\sum_{k=1}^N a_k h[n-k] + \sum_{k=0}^M b_k \delta[n-k], \quad n \geq 0 \quad (22)$$

sau, echivalent,

$$h[n] = -\sum_{k=1}^N a_k h[n-k] + b_n, \quad 0 \leq n \leq M \quad (23)$$

iar pentru $n > M$ relația (22) se reduce la:

$$h[n] = -\sum_{k=1}^N a_k h[n-k], \quad n > M \quad (24)$$

Dacă $H_d(z)$ este un filtru cu poli și zerouri, răspunsul său la $\delta[n]$ ar trebui să satisfacă aceleași ecuații (22) – (24), lucru care, în general, nu se întâmplă. Pe baza relației (24) se definește valoarea de predicție liniară a lui $h_d[n]$, ca fiind

$$\hat{h}_d[n] = -\sum_{k=1}^N a_k h_d[n-k], \quad n > M \quad (25)$$

Ca și în cazul filtrului numai cu poli, suma pătratelor erorii de predicție este:

$$\varepsilon_1 = \sum_{n=M+1}^{\infty} (h_d[n] - \hat{h}_d[n])^2 = \sum_{n=M+1}^{\infty} \left(h_d[n] + \sum_{k=1}^N a_k h_d[n-k] \right)^2 \quad (26)$$

Minimizarea lui ε_1 în raport cu coeficienții $\{a_k\}$ conduce la setul de ecuații liniare:

$$\sum_{l=1}^N a_l r_{dd}[k, l] = -r_{dd}[k, 0], \quad k = 1, 2, \dots, N \quad (27)$$

$$\text{unde } r_{dd}[k, l] = \sum_{n=M+1}^{\infty} h_d[n-k] h_d[n-l] \quad (28)$$

Ecuațiile (27), care dau coeficienții $\{a_k\}$, se reduc la aproximarea filtrului numai cu poli, când M se impune a fi zero.

Parametrii $\{b_k\}$ ce determină zerourile filtrului se obțin din relația (23), unde $h[n] = h_d[n]$, prin înlocuirea valorilor $\{\hat{a}_k\}$ obținute din (27).

$$b_n = h_d[n] + \sum_{k=1}^N \hat{a}_k h_d[n-k], \quad 0 \leq n \leq M \quad (29)$$

În concluzie, parametrii $\{\hat{a}_k\}$ ce determină polii se obțin prin metoda celor mai mici pătrate, în timp ce parametrii $\{b_k\}$, care determină zerourile, se obțin ca în metoda de aproximare Padé. Această tehnică pentru determinarea polilor și zerourilor lui $H(z)$ se numește **metoda Prony**.

Metoda celor mai mici pătrate furnizează estimați buni pentru parametrii polilor. Metoda Prony poate să nu fie eficace pentru estimarea parametrilor $\{b_k\}$, în primul rând deoarece calculul lui b_k din (29) nu se bazează pe o metodă ce minimizează un criteriu de eroare.

O altă metodă, în care ambele seturi de parametri, $\{a_k\}$ și $\{b_k\}$, se determină pe baza celor mai mici pătrate a fost propusă de **Shanks** (1967). Parametrii $\{a_k\}$ sunt calculați prin aplicarea criteriului celor mai mici pătrate ca în relația (27), prezentată anterior. Aceasta produce estimății $\{\hat{a}_k\}$ ce permit sintetizarea filtrului numai cu poli:

$$H_1(z) = \frac{1}{1 + \sum_{k=1}^N \hat{a}_k z^{-k}} \quad (30)$$

Răspunsul acestui filtru la $\delta[n]$ este:

$$v[n] = -\sum_{k=1}^N \hat{a}_k v[n-k] + \delta[n], \quad n \geq 0 \quad (31)$$

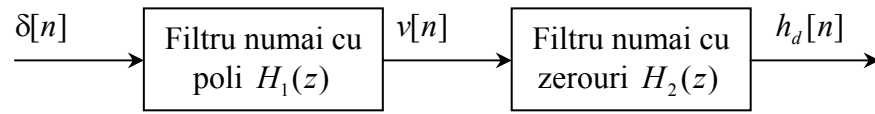


Figura 3: Metoda celor mai mici pătrate pentru determinarea polilor și zerourilor unui filtru

Dacă secvența $\{v[n]\}$ este folosită pentru a excita un filtru numai cu zerouri cu funcția de sistem

$$H_2(z) = \sum_{k=0}^M b_k z^{-k} \quad (32)$$

după cum se arată în figura 3, răspunsul acestuia este

$$\hat{h}_d[n] = \sum_{k=0}^M b_k v[n-k] \quad (33)$$

Se poate defini o secvență de eroare

$$e[n] = h_d[n] - \hat{h}_d[n] = h_d[n] - \sum_{k=0}^M b_k v[n-k] \quad (34)$$

și, în consecință, parametrii $\{b_k\}$ pot fi de asemenea determinați cu ajutorul criteriului celor mai mici pătrate și anume, prin minimizarea relației

$$\varepsilon_2 = \sum_{n=0}^{\infty} \left(h_d[n] - \sum_{k=0}^M b_k v[n-k] \right)^2 \quad (35)$$

Astfel, se obține un set de ecuații liniare pentru parametrii $\{b_k\}$, în forma

$$\sum_{k=0}^M b_k r_{vv}[k, l] = r_{hv}[l], \quad l = 0, 1, \dots, M \quad (36)$$

unde, prin definiție

$$r_{vv}[k, l] = \sum_{n=0}^{\infty} v[n-k]v[n-l] \quad (37)$$

$$r_{hv}[l] = \sum_{n=0}^{\infty} h_d[n]v[n-l] \quad (38)$$

1.3 Metoda directă de sinteză a filtrelor IIR Yule-Walker

Există însă și tehnica de proiectare a filtrelor IIR în care optimizarea parametrilor filtrului se face în domeniul frecvență prin metoda celor mai mici pătrate. Pentru această metodă se va folosi funcția:

[b, a]=yulewalk(n, f, m); ce va returna coeficienții unui filtru recursiv de ordinul n. Parametrii de intrare sunt:

-n ordinul filtrului IIR;

-f și m sunt vectorii ce conțin frecvențele respectiv răspunsurile în amplitudine la frecvențele specificate în f. Frecvențele din vectorul f trebuie să se încadreze în intervalul [0.0 1.0]. Cu **plot(f, m)** se poate vizualiza răspunsul în frecvență dorit.

Exemplu:

```
%program P9_0
f=[0 0.3 0.4 0.7 0.8 1];
m=[1 1 0.01 0.01 0.5 0.5];
[b,a]=yulewalk(10,f,m);
[h,w]=freqz(b,a);
plot(w/pi, abs(h));
```

2. Desfășurarea lucrării

Se introduc următoarele funcții Matlab:

[b, a]=pademethod(hd, M, N); returnează coeficienții unui filtru digital, [b, a], sintetizat cu ajutorul metodei Padé. Funcția are ca argumente de intrare răspunsul la impuls dorit hd și numărul de zerouri, respectiv poli ai filtrului sintetizat, M, N;

a=lsinv(hd, N); returnează coeficienții de la numitor ai unui filtru doar cu poli sintetizat prin metoda celor mai mici pătrate. Parametrii de intrare sunt: răspunsul la impuls dorit, respectiv numărul de poli ai filtrului;

[b, a]=prony(hd, M, N); returnează coeficienții unui filtru digital, sintetizat cu ajutorul metodei Prony.

[b, a]=shanksmet(hd, M, N); returnează coeficienții unui filtru digital, sintetizat cu ajutorul metodei Prony.

Exemplul 1. Se consideră filtrul Butterworth de ordin 4 cu funcția sistem

$$H_d(z) = \frac{0.0048(z+1)^4}{(z^2 - 1.0486z + 0.2962)(z^2 - 1.3210z + 0.6328)}$$

Având răspunsul la impuls $h_d[n]$ al filtrului de mai sus, să se utilizeze *metoda de aproximare Padé* pentru a proiecta filtrul IIR a cărui răspuns la impuls $h[n]$ aproximează cel mai bine pe $h_d[n]$.

Soluție:

Se observă că filtrul dorit are $M = 4$ zerouri și $N = 4$ poli. Este interesant de determinat coeficienții prin aproximarea Padé atunci când numărul de zerouri și/sau poli nu este identic cu numărul parametrilor filtrului dorit. Se scriu ecuațiile de determinare a coeficienților $\{a_k\}$ și $\{b_k\}$. Folosind funcția creată **pademethod.m** se rezolvă sistemul de ecuații, determinându-se coeficienții filtrului în patru cazuri: $(M = 3, N = 5)$; $(M = 3, N = 4)$; $(M = 4, N = 4)$; $(M = 4, N = 5)$. Se observă că pentru $M = 3$, răspunsul în frecvență corespunzător este o aproximare slabă a răspunsului dorit. În acest caz, o creștere a lui N de la 4 la 5 nu îmbunătățește răspunsul în frecvență al sistemului. (se va utiliza scriptul **P9_1.m**).

Exemplul 2. Se consideră același filtru de la exemplul 1 și se aproximează filtrul cu ajutorul *metodei de proiectare a filtrului invers (numai cu poli) pe baza celor mai mici pătrate*. Se rulează scriptul **P9_2.m** pentru mai multe valori ale ordinului filtrului. ($N=4, N=9, N=20$). Ce se observă? De asemenea schimbați lungimea răspunsului la impuls ($L=21, L=100$). Se observă vreo influență a lungimii răspunsului la impuls asupra răspunsului filtrului sintetizat?

Exemplul 3. Cu aceleași cerințe de la exemplul 1, se proiectează un filtru IIR prin metoda Prony. (**P9_3.m**). Ce se observă atunci când modificați lungimea răspunsului la impuls L ?

Exemplul 4. Cu aceleași cerințe de la exemplul 1, se proiectează un filtru IIR prin metoda Shanks. (**P9_4.m**).

3. Aplicații propuse

1. Fie filtrul trece-jos Chebyshev de tip II cu 3 poli și 3 zerouri ce are funcția sistem:

$$H_d(z) = \frac{0.0736 + 0.0762z^{-1} + 0.0762z^{-2} + 0.0736z^{-3}}{1 - 1.3969z^{-1} + 0.8778z^{-2} - 0.1812z^{-3}}$$

Utilizați *metoda de aproximare Padé* a lui $H_d(z)$ și comparați performanțele metodei în cazul în care numărul de zerouri și poli ai filtrului proiectat variază ($M = 2, N = 3$; $M = 2, N = 4$; $M = 3, N = 3$; $M = 3, N = 4$).

2. Aproximați FTJ Chebyshev de tip II din aplicația propusă 1, folosind cele trei metode bazate pe cele mai mici pătrate (*metoda de proiectare a filtrului pe baza celor mai mici pătrate, metoda Prony și metoda Shanks*).

3. Sintetizați un filtru trece bandă cu ajutorul metodei Yule-Walker care are banda de trecere între frecvențele $[0.4 \ 0.7]$, iar frecvențele 0.3 și 0.8 delimitează cele două benzi de oprire.