

Modelovanie sezónnosti úspor domácností stavovými modelmi

Dušan MARČEK*

Úvod

Prevažná časť ekonometrického modelovania sa zaoberá opisom makroekonomických procesov. Kvantifikácia týchto modelov je založená na dátach s ročnou periodicitou. Je to pochopiteľné, lebo kvantifikácia modelov z dát s ročnou periodicitou je jednoduchšia, nevyžaduje zohľadňovať sezónnosť. Sezónnosť je typická pre ekonomické procesy, ktoré prebiehajú, a dáta o nich evidované sú s periodicitou menšou než jeden rok. Vyrovnáť sa so sezónnosťou v ekonometrických modeloch predstavuje zložitý problém. Jedno riešenie v tomto vyrovnaní so sezónnosťou v modelovaní finančných trhov, napríklad pri modelovaní volatility finančných aktív s využitím vysokofrekvenčných dát, je naznačené v [3]. Celý proces modelovania sa transformuje do inej sústavy s premenlivou (z hustenou alebo rozšírenou) postupnosťou časových periód, tzv. operačného času. Všetky výpočty spojené s kvantifikáciou a so štatistickým overovaním modelu sa vykonávajú na báze nového operačného času. Transformáciou fyzického času do operačného času sa síce odstraňujú sezónne efekty v rámci jednotlivých periód, ale vznikajú nové, nežiaduce efekty týkajúce sa porušenia štandardných predpokladov o náhodných zložkách rovníc modelu. Ide v podstate o riešenie, v ktorom sa fakticky eliminuje, odfiltruje, a teda nemodeluje sezónnosť časového radu.

Ďalšia možnosť transformovania časového radu zaťaženého sezónnosťou je vytvorenie časového radu kľzavých súčtov. V ekonometrických modeloch sa táto možnosť nedá prakticky využiť. Vytváraním časových radov kľzavých súčtov sa jednak znižuje počet pozorovaní, ale hlavne sa deformujú závislosti v štruktúrnych modeloch. Prakticky je znemožnená jeho analytická špecifikácia. Iná, klasická možnosť zohľadnenia sezónnosti, ktorú poskytuje štatistika, je priame zakomponovanie sezónnej zložky (konštantnej alebo proporčionálnej k trendu) do ekonometrického modelu. V tomto riešení nastávajú problémy s kvantifikáciou ekonometrických modelov, predovšetkým ak ide o viacrovnicový model. Z tohto dôvodu uvedené modely nenašli v praxi širšie uplatnenie.

* prof. Ing. Dušan MARČEK, CSc., Žilinská univerzita, Fakulta riadenia a informatiky, Katedra makro a mikroekonomiky, Moyzesova 20, 010 26 Žilina

Na riešenie problému sezónnosti v ekonometrickom modelovaní, t. j. pri špecifikácii premenných, navrhovaní funkčných tvarov a pri jeho riešení, t. j. pri odhade hodnôt endogénnych premenných, možno využiť Kalmanov filter. Kalmanovou filtráciou sa rozumie identifikačný prístup, ktorý sa používa na zisťovanie stavových premenných procesov [1]. Ukážeme, že odhady hodnôt veličín modelu možno získať pomocou Kalmanových rekurzií. V našich podmienkach je na odhad hodnôt veličín ekonometrického modelu známa práca [5], v ktorej sa opisuje proces vývoja inflácie. Aplikácia Kalmanových rekurzií je podmienená zápisom modelu v tzv. stavovej forme jednotlivých štruktúrnych rovníc modelu.

V rámci príspevku postupne rozoberieme definíciu všeobecnej stavovej formy štruktúrnych modelov. Ukážeme, ako je možné zakomponovať do modelu sezónnosť veličín. V druhej časti uvedieme formuly Kalmanových rekurzií. V tretej časti sa budeme zaoberať predikciami stavových modelov, vo štvrtej časti ilustrujeme Kalmanove rekurzie na odhad úspor obyvateľstva, a nakoniec porovnáme presnosť týchto odhadov s klasickým štatistickým prístupom.

1. Všeobecná reprezentácia štruktúrnych modelov v stavovej forme – modely so sezónnou zložkou

V ďalšej časti sa budeme zaoberať modelmi diskretných časových radov, ktoré možno vyjadriť v tzv. lineárnej stavovej forme. Rozvinutá reprezentácia stavovej formy modelu je opísaná vzťahmi v tvare

$$Y_t = G_t X_t + W_t \quad (2)$$

$$X_t = F_t X_{t-1} + V_t \quad (3)$$

kde Y_t je vektor alebo skalár tvorený pozorovaniami výstupných premenných (výstupu) skúmaného procesu; $\{Y_t\}$ je závislý od iného procesu $\{X_t\}$; X_t je stavový vektor systému v čase t . Lineárny proces $\{X_t\}$ je opísaný rovnicou (2). F_t a G_t sú známe časové postupnosti matic alebo vektorov, ktoré môžu, ale nemusia závisieť od času t . Predpokladá sa, že prvky matice F dimenzie ($v \times v$) a matice G dimenzie ($w \times v$), kde v označuje počet stavových premenných, w označuje počet výstupných premenných, sú známe. Získajú sa spravidla na základe matematicko-štatistickej analýzy procesu. W_t , V_t sú nezávislé gaussovské náhodné veličiny s nulovou strednou hodnotou a so známymi kovariančnými maticami $E(W_t W_t') = R_t$, $E(V_t V_t') = Q_t$, kde $\{R_t\}$ je postupnosť matic dimenzie ($w \times w$), $\{Q_t\}$ je postupnosť matic dimenzie ($v \times v$), pre $t = 1, 2, \dots$. Rovnicu (3) možno ďalej zovšeobecniť tak, aby na jej pravej strane bol zahrnutý člen $H_t u_t$, kde H_t je daná postupnosť matic parametrov exogénnych premenných, u_t je vektor špecifikovaných exogénnych vstupných premenných, ktoré ovplyvňujú proces $\{X_t\}$

v čase $t = 1, 2, \dots$ V tomto prípade vývoj stavovej veličiny \mathbf{X}_t – vektora dimenzie $(v \times 1)$ systému v čase t namiesto tvaru (3) bude mať tvar

$$\mathbf{X}_t = \mathbf{F}_t \mathbf{X}_{t-1} + \mathbf{H}_t \mathbf{u}_t + \mathbf{V}_t$$

V ďalšom texte budeme kvôli jednoduchosti predpokladať, že matice, \mathbf{F}_t , \mathbf{G}_t , \mathbf{R}_t , \mathbf{Q}_t , nebudú závisieť od času t .

Z rovnice (2) vyplýva, že pozorovania procesu $\{\mathbf{Y}_t\}$ sú funkciou, \mathbf{X}_0 , \mathbf{V}_1 , \mathbf{V}_2 , ..., \mathbf{V}_{t-1} a \mathcal{W}_t , t. j.

$$\mathbf{Y}_t = g_t(\mathbf{W}_0, \mathbf{V}_1, \mathbf{V}_2, \dots, \mathbf{V}_{t-1}, \mathcal{W}_t)$$

a vývoj stavovej veličiny $\{\mathbf{X}_t\}$ je funkciou \mathbf{X}_0 , $\mathbf{V}_1, \mathbf{V}_2, \dots, \mathbf{V}_{t-1}$, t. j.

$$\mathbf{Y}_t = f_t(\mathbf{X}_0, \mathbf{V}_1, \mathbf{V}_2, \dots, \mathbf{V}_{t-1})$$

Reprezentácia stavovej formy modelu podľa rovníc (2) a (3) je veľmi jednoduchá. Podľa nej napríklad vývoj $\{\mathbf{Y}_t\}$ je determinovaný procesom $\{\mathbf{X}_t\}$ opísaným rovnicou (2). Štruktúrny tvar modelu časového radu umožňuje explicitné zohľadnenie sezónnej zložky modelov časových radov ekonomických veličín. V ďalšej časti ukážeme na príkladoch zápisy niektorých jednoduchých štruktúrnych modelov ekonomických procesov so sezónnou zložkou v reprezentácii stavovej formy.

Reprezentantom týchto modelov môže byť model lokálne lineárneho trendu so sezónnou zložkou, buď s konštantnou sezónnosťou, alebo s proporcionálnou sezónnosťou. Ukážeme ich východiskové (formulačné) tvary a tvary v stavovej forme, čím priblížime ich význam a ukážeme použitie zavedenej symboliky.

Model lokálne lineárneho trendu s konštantnou sezónnosťou

Jeho jeden možný tvar je

$$y_t = x_{t,1} + x_{t,2} \quad (4)$$

$$x_{t,1} = x_{t-1,1} + v_t \quad (5)$$

$$x_{t,2} = \sum_{j=1}^{s-1} x_{t-j,2} \quad (6)$$

V prvej rovnici (4) je výstupná premenná y_t modelovaná súčtom trendovej zložky (stavová premenná $x_{t,1}$) a sezónnej zložky s konštantnou sezónnosťou $x_{t,2}$. Obidve stavové premenné sú opísané vlastným vývojovým modelom. Lokálny konštantný trend (druhá rovnica) je stochastický proces s náhodnou zložkou v_t . O hodnote rozptylu náhodnej zložky v_t nemáme žiadne apriórne informácie. Budeme ju považovať za jeden z parametrov stavovej reprezentácie modelu, ktorý treba štatisticky odhadnúť. O sezónnom vplyve sa predpokladá, že pôsobí v rámci roka, t. j. v rámci s sezón s kompezačným účinkom v rámci s sezón. Rovnica (6) vyjadruje tento predpoklad (pozri [2]).

Model má spolu s stavových premenných $x_{t, 1}, x_{t, 2}, x_{t-1, 2}, x_{t-2, 2}, \dots, x_{t-s+1, 2}$ a jednu výstupnú premennú y_t , z čoho vyplýva $v = s, w = 1$. Môžeme stanoviť rozmery matíc $\mathbf{X}, \mathbf{Y}, \mathbf{F}, \mathbf{G}, \mathbf{R}$ a \mathbf{Q} na konštrukciu lineárneho trendu so sezónnou zložkou v stavovej forme. Možný tvar modelu v stavovej forme (2) a (3) pre $s = 4$ je

$$\mathbf{Y}_t = \mathbf{G} \mathbf{X}_t + \mathbf{W}_t$$

$$y_t = [1 \ 1 \ 0 \ 0] \begin{bmatrix} x_{t, 1} \\ x_{t, 2} \\ x_{t-1, 2} \\ x_{t-2, 2} \end{bmatrix} + 0$$

$$\mathbf{X}_t = \mathbf{F} \mathbf{X}_{t-1} + \mathbf{V}_t$$

$$\begin{bmatrix} x_{t, 1} \\ x_{t, 2} \\ x_{t-1, 2} \\ x_{t-2, 2} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & -1 & -1 \\ 0 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_{t-1, 1} \\ x_{t-1, 2} \\ x_{t-2, 2} \\ x_{t-3, 2} \end{bmatrix} + \begin{bmatrix} v_t \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\mathbf{Q} = \begin{bmatrix} q_1 & & & \\ & 0 & 0 & \\ & 0 & 0 & \\ & & & 0 \end{bmatrix}; \mathbf{R} = \sigma_{\mathcal{E}}^2 = 0, \text{ kde } q_1 = \sigma_v^2$$

Všetky uvedené štruktúrne modely majú parametre reprezentované prvkami prenosovej matice \mathbf{F} a matice \mathbf{G} rovné 1, -1 alebo 0. Sú známe priamo už pri formulácii štruktúrnych modelov. Všeobecne prvky matice \mathbf{F} vyjadrujú intenzitu vzťahov medzi jednotlivými stavovými veličinami, medzi stavovými veličinami a výstupnou veličinou, príp. exogénnymi veličinami. Ich hodnoty sa spravidla odhadujú pomocou vhodných štatistických, resp. ekonometrických metód.

Model lokálne lineárneho trendu s proporcionálnou sezónnosťou

Štruktúrny tvar modelu je obdobou modelu s konštantnou sezónnosťou, s tým rozdielom, že jeho sezónna zložka je vyjadrená proporcionálne k úrovni trendovej zložky v každom pozorovaní. Jeho tvar je

$$y_t = x_{t, 1} + x_{t, 2} / x_{t, 1} \quad (7)$$

$$x_{t, 1} = x_{t-1, 1} + v_t \quad (8)$$

$$x_{t,2}/x_{t,1} = -\sum_{j=1}^{s-1} x_{t-j,2}/x_{t-j,1} \quad (9)$$

Čitateľ sa môže analógiou ľahko presvedčiť, že v modeloch (7) až (9) v ich štruktúrnej stavovej forme pre $s = 4$ budú jednotlivé matice mať tvar

$$\mathbf{X}_t = [x_{t,1}, x_{t,2}/x_{t,1}, x_{t-1,2}/x_{t-1,1}, x_{t-2,2}/x_{t-2,1}]'; \quad \mathbf{G} = [1 \ 1 \ 0 \ 0]$$

$$\mathbf{X}_{t-1} = [x_{t-1,1}, x_{t-1,2}/x_{t-1,1}, x_{t-2,2}/x_{t-2,1}, x_{t-3,2}/x_{t-3,1}]'; \quad q_1 = \sigma_v^2; \quad \mathbf{R} = 0$$

$$\mathbf{Q} = \begin{bmatrix} q_1 & & & \\ & 0 & 0 & \\ & 0 & 0 & \\ & & & 0 \end{bmatrix}; \quad \mathbf{F} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & -1 & -1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

2. Kalmanove rekurzcie a počiatočné hodnoty Kalmanovho filtra

Základnou úlohou Kalmanových rekurzii, ktoré sú spojené so stavovými formami modelov (2), (3), je určenie najlepších – v zmysle minimálnej strednej štvorcovej chyby – lineárnych odhadov pozorovaných hodnôt stavového vektora \mathbf{X}_t na báze dostupných informácií z pozorovaní výstupnej veličiny Y_t . Riešenia pre odhady stavového vektora \mathbf{X}_t na základe dostupnej informačnej množiny I_t obsahujúcej pozorovania veličiny Y_t až do periódy t sa nazývajú *Kalmanove rekurzcie*. Kalmanovými rekurziami sa rieši: (a) filtračný problém (Kalmanova filtrácia), (b) problém vyrovnávania hodnôt stavového vektora \mathbf{X}_t (*smoothing problem* – Kalmanove vyrovnávanie), (c) Kalmanova predikcia. V súlade s [3] riešenie uvedených troch problémov Kalmanovými rekurziami ďalej opíšeme.

Kalmanova filtrácia. Kalmanovou filtráciou sa rieši odhad hodnôt stavového vektora \mathbf{X}_t a strednej kvadratickej chyby odhadu na základe informačnej množiny. Označme čas pozorovania písmenom r , najlepší lineárny estimátor hodnôt stavového vektora \mathbf{X}_t založeného na informačnej množine pozorovaní endogénnej veličiny až do času a vrátane času r . Označme ďalej strednú hodnotu štvorcovej chyby odhadu (kovariančnú maticu) $P_{t|r} = E\left[\left(X_t - \hat{X}_{t|r}\right)\left(X_t - \hat{X}_{t|r}\right)'\right]$.

Potom odhady hodnôt stavového vektora a ich kovariančných matic sú dané

$$\hat{X}_{t|r} = \hat{X}_{t|t-1} + K_t \left(Y_t - G_t \hat{X}_{t|t-1} \right) \quad (10)$$

$$\hat{X}_{t+1|t} = F_{t+1} \hat{X}_{t|r} \quad (11)$$

$$P_{t|r} = P_{t|t-1} - K_t G_t P_{t|t-1} \quad (12)$$

$$P_{t+1|t} = Q_{t+1} + F_{t+1} P_{t|t} F'_{t+1} \quad (13)$$

kde

$$K_t = P_{t|t-1} G'_t (G_t P_{t|t-1} G' + R_t)^{-1} \quad (14)$$

je tzv. matica zosilnenia. Rovnicami (10) až (13) sa odhadujú hodnoty stavového vektora $X_{t|t}$ a kovariančnou maticou chyby odhadu $P_{t|t}$ rekurzívne, až matica K_t konverguje ku konštantnej hodnote. Vidíme, že pre Kalmanovu filtráciu musia byť známe matice G , F modelu (2), (3), počiatkový odhad matic $P_{t|t}$, R , Q a dáta výstupu y_t .

Kalmanove vyrovnávanie. Kalmanov filter poskytuje odhady hodnôt stavovej veličiny $X_{t|t}$ v čase t na základe informácií z pozorovaní veličiny Y_t až do času t . Len čo sú k dispozícii ďalšie (dodatočné) informácie, t. j. $\{Y_{t+i}\}$, $i=1, 2, \dots, r$, odhad X_t možno zlepšiť zohľadnením týchto dodatočných informácií. Kalmanove vyrovnávanie poskytuje procedúry na získanie odhadu X_t na základe informačnej množiny I_T , kde $T > t$. Je dané nasledujúcimi rekurentnými vzťahmi

$$\hat{X}_{t|T} = \hat{X}_{t|t} + J_t (X_{t+1|T} - \hat{X}_{t+1|t}) \quad (15)$$

$$P_{t|T} = P_{t|t} + J_t (P_{t+1|T} - P_{t+1|t}) J'_t \quad (16)$$

kde

$$J_t = P_{t|t} F'_{t+1} P_{t+1|t}^{-1} \quad (17)$$

Na to, aby sa získali vyrovnané hodnoty $\hat{X}_{t|T}$ a $P_{t|T}$ rekurziami (15) a (16), musí sa najskôr vykonať Kalmanova filtrácia na získanie hodnôt $\hat{X}_{t|t}$ a $P_{t|t}$, čo vyžaduje zadanie počiatkových hodnôt matic G , F , P_0 , R , Q a dáta y_0 .

Kalmanova predikcia. Kalmanovými predikčnými procedúrami sa vypočítajú budúce hodnoty stavového vektora, t. j. vypočíta sa $X_{t+\tau}$, $\tau > 0$, založené na aktuálnej (súčasnej) informačnej množine I_t . Odhad týchto hodnôt je daný

$$\hat{X}_{t+\tau|t} = F_{t+\tau} \hat{X}_{t+\tau-1|t} \quad (18)$$

a kovariančná matica

$$P_{t+\tau|t} = F_{t+\tau} P_{t+\tau-1|t} F'_{t+\tau} + Q_{t+\tau} \quad (19)$$

Na získanie predikčných hodnôt $\{\hat{X}_{t+\tau|t}, P_{t+\tau|t}\}$ rekurzívne výrazy (18) a (19) začínajú výpočet s hodnotami $\{X_{t+1|t}, P_{t+1|t}\}$, ktoré sa získajú z Kalmanovej filtrácie. Ak sú vypočítané predikčné hodnoty $\{X_{t+\tau|t}\}$, potom z výrazu (1) je možné získať aj predikčné hodnoty $\hat{Y}_{t+\tau|t}$ ako

$$\hat{Y}_{t+\tau|t} = G_{t+\tau} \hat{X}_{t+\tau|t} \quad (20)$$

a variančnú maticu $\Sigma_{t+\tau|t}$ stredných kvadratických chýb týchto predpovedí

$$\Sigma_{t+\tau|t} = E \left[\left(Y_{t+\tau|t} - \hat{Y}_{t+\tau|t} \right) \left(Y_{t+\tau|t} - \hat{Y}_{t+\tau|t} \right)' \right] = G_{t+\tau} P_{t+\tau|t} G_{t+\tau}' + R_{t+\tau} \quad (21)$$

Všetky Kalmanove rekurzie, t. j. filtrácia, vyrovnávanie a predikcia využívajú rekurzívne procedúry Kalmanovho filtra, dané rovnicami (10) až (13). Rovnica (10) Kalmanovho filtra vyžaduje znalosť počiatočných hodnôt stavového vektora $\hat{X}_{t|t-1}$ a kovariančnej matice $P_{t|t-1}$. Na riešenie týchto počiatočných hodnôt bolo navrhnutých niekoľko spôsobov. Jeden zo spôsobov využíva informácie z pozorovaní výstupnej premennej Y_t . Stavový vektor \mathbf{X}_t je stĺpcový vektor obsahujúci v prvkov. Na odhad počiatočných hodnôt všetkých v prvkov stavového vektora $\mathbf{X}_{v+1|v}$ a na inicializáciu matice $P_{v+1|v}$ sa môže použiť prvých v členov zo vstupu Y_t a potom hodnoty $\mathbf{X}_{v+1|v}$ a $P_{v+1|v}$ vypočítať z Kalmanovej filtrácie.

3. Odhad rozptylov štruktúrneho modelu a predikcia

Na výpočet budúcich hodnôt stavových premenných $\mathbf{X}_{t+\tau}$ a na aplikáciu Kalmanovho filtra vôbec, ako to vyplýva z výrazov pre Kalmanove rekurzie, t. j. aby sme ich mohli vykonať pomocou rekurzívnych Kalmanových procedúr, je nevyhnutná znalosť počiatočných hodnôt stavového vektora $\hat{\mathbf{X}}_{1|0}$ a kovariančnej matice $P_{1|0}$. V predchádzajúcej časti sme uviedli, že ich počiatočný odhad je možné vykonať Kalmanovým filtrom z prvých v hodnôt výstupu Y_t . Na odhad počiatočných hodnôt matice $P_{1|0}$ je nevyhnutné numericky špecifikovať matice \mathbf{Q} a \mathbf{R} , ktoré obsahujú v hlavnej diagonále rozptyly náhodných zložiek stavového modelu. Počiatočný odhad kovariančnej matice $P_{1|0}$ je preto aj funkciou rozptylov vystupujúcich v maticiach \mathbf{Q} a \mathbf{R} . Aby bol problém počiatočného odhadu matice $P_{1|0}$ možný, musí byť ešte predtým vyriešená otázka odhadu numerických hodnôt rozptylov náhodných zložiek stavového modelu.

Jedno možné riešenie odhadu rozptylov náhodných zložiek stavového modelu je založené na predpoklade, že existuje východisková predstava o ich hodnotách, t. j. rozptyly sú a priori známe. Riešenie je vhodné v simulovaných technických procesoch, na základe ktorých chceme skúmať správanie určitej premennej v rôznych situáciách. V ďalšej časti sa však sústredíme na spôsob odhadu rozptylov stavového modelu z dát výstupu Y_t tak, aby tieto odhady zaručovali maximálnu vierohodnosť stavového modelu.

Uvažujme stavový model daný rovnicami (2) a (3) s parametrami danými maticami \mathbf{F} , \mathbf{G} , \mathbf{Q} , \mathbf{R} . Ako sme už uviedli a na príkladoch ilustrovali, parametre, ktoré stoja pri stavových premenných, musia byť vopred známe. Matica \mathbf{Q} je

symetrická dimenzia ($v \times v$). Na hlavnej diagonále obsahuje hodnoty rozptylov náhodných zložiek stavových rovníc. Na zjednodušenie označme tieto prvky ako q_1, q_2, \dots . Napríklad v modeli danom rovnicami (4) až (6) bude: $q = \sigma_v^2$. Analogicky označme rozptyl náhodnej zložky pozorovaného procesu $\sigma_\varepsilon^2 = r_1 = 0$. Vidíme, že štruktúrny model je kompletne parametrizovaný parametrami matíc F, G, Q, R . Všetky ich komponenty zhromaždíme do vektora, ktorý označíme θ .

Funkciu vierohodnosti pozorovaní $\{Y_1, Y_2, \dots, Y_N\}$ je možné jednoducho napísať ako

$$L(\theta; Y_1, Y_2, \dots, Y_N) = \prod_{t=1}^N N(Y_{t|t-1}, \Sigma_{t|t-1}) \quad (22)$$

kde

$$\Sigma_{t|t-1} = E \left[\left(Y_t - \hat{Y}_{t|t-1} \right) \left(Y_t - \hat{Y}_{t|t-1} \right)' \right]$$

Logaritmus funkcie vierohodnosti (22) je

$$\ln L(\theta; Y_1, Y_2, \dots, Y_N) = -\frac{1}{2} \sum_{t=1}^N \ln \left| \Sigma_{t|t-1} \right| - \frac{1}{2} \left(Y_t - \hat{Y}_{t|t-1} \right)' \Sigma_{t|t-1}^{-1} \left(Y_t - \hat{Y}_{t|t-1} \right) + konst \quad (23)$$

Vo výraze pre logaritmus funkcie vierohodnosti (23) sú sumačné limity od 1 do N . Ak sa prvých v hodnôt použije na inicializáciu premenných pre Kalmanove rekurzcie, dolné limity sumácií v (23) musia byť upravené na hodnotu $t = v + 1$. Potom $\ln L$ je podmienený $\ln L$, a to podmienený fixovanými hodnotami vstupov Y_1, y_2, \dots, y_v .

Postavením prvých derivácií zápornej funkcie (23) podľa q_1, q_2, \dots a r_1 rovných nule sa získajú rovnice sústavy, z ktorých možno určiť maximálne vierohodné estimátory rozptylov $\hat{q}_1, \hat{q}_2, \dots$ a \hat{r}_1 .

Ako sme uviedli v predchádzajúcich častiach, odhadová metóda ML je citlivá na počiatočné hodnoty, čo sa prejavuje v konvergencii alebo nekonvergencii metódy, resp. vo výpočtovej zložitosti a rýchlosti odhadu. Úlohy tohto typu sa nedajú reálne riešiť bez použitia počítača a vhodného štatistického softvéru.

Ak sú vypočítané hodnoty rozptylov q_1, q_2, \dots a r_1 , môžu sa na základe nich spresniť odhady počiatočných hodnôt stavového vektora $\hat{X}_{t|t-1}$ a kovariančnej matice $P_{1|0}$ a potom prostredníctvom nich spresniť výstupy aplikovaním Kalmanových rekurzii.

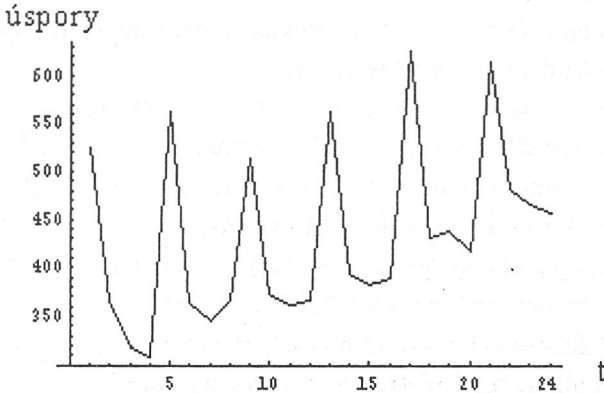
4. Modelovanie úspor obyvateľstva

Na ilustráciu aplikácie stavového modelu a Kalmanovej filtrácie v ekonomicke uvažujme časový rad štvrťročných dát za roky 1961–1966 úspor obyvateľstva v Štátnej sporiteľni, vyjadrených v 10 mil. Kčs [2]. Označme tieto pozorovania

časového radu symbolom y_t pre $t = 1, 2, \dots, N$, kde N je celkový počet pozorovaní a $N = 24$. Vývoj časového radu úspor obyvateľstva y_t je graficky znázornený na obrázku 1.

Obrázok 1

Vývoj úspor obyvateľstva y_t ako funkcia času



Z obrázku 1 je vidieť, že časový rad $\{y_t\}$ vykazuje v jednotlivých rokoch systematické výkyvy, pričom v jednotlivých rokoch má vzostupnú tendenciu. Z týchto zistení sa rozhodneme, že časový rad budeme modelovať lokálne lineárnym trendom so sezónnou zložkou v tvare rovníc (4) až (6).

O hodnote rozptylu náhodnej zložky v_t nemáme žiadne apriórne informácie. Budeme ju považovať za jeden z parametrov stavového modelu, ktorý treba štatisticky odhadnúť.

Ak je model napísaný v stavovej forme so špecifikovanými prvkami prenosovej matice F , pomocou Kalmanových rekurzií môžeme odvodiť (generovať) hodnoty stavových premenných lokálne lineárneho trendu $x_{t,1}$ a sezónnej zložky $x_{t,2}$, prípadne generovať ich budúce hodnoty.

Proces generovania hodnôt stavových premenných, príp. predikcia ich hodnôt sa vykonáva postupne procedúrami, ktoré sme pomenovali ako Kalmanova filtrácia, vyrovnávanie a predikcia. Predtým než prikrôčime k vykonaniu týchto procedúr, musíme určiť počiatočné hodnoty stavového vektora $\hat{X}_{1|0}$ a kovariančnej matice $P_{1|0}$.

Počiatočné hodnoty vektora $\hat{X}_{1|0}$ a matice $P_{1|0}$ určíme z prvých $v = 4$ hodnôt pozorovaní výstupu Kalmanovou filtráciou pomocou výrazov (10) až (14) pre $\hat{X}_{v+1|v}$. Odhad počiatočných hodnôt vektora $\hat{X}_{v+1|v} = [378.75 \ 145.25 \ -69.75 \ -61.75]'$ a matice $P_{v+1|v}$ je:

$$P_{v+1|v} = q_1 \begin{bmatrix} 1.875 & 0.0625 & -0.875 & -0.125 \\ 0.0625 & 0.875 & -0.625 & -0.375 \\ -0.875 & -0.625 & 0.875 & 0.125 \\ -0.125 & -0.375 & 0.125 & 0.375 \end{bmatrix}$$

Vidíme, že počiatočné hodnoty $P_{v+1|v}$ sú funkciou neznámeho parametra q_1 . Jeho odhad nájdeme minimalizáciou funkcie logaritmu vierohodnosti (23) pre parameter q_1 , splňajúci vzťahy dané maticou $P_{v+1|v}$.

Aplikácia programu pre hľadanie minima záporného logaritmu podmienenej vierohodnosti pre uvedené podmienky dáva riešenie $q_1 = 701.925$.

Štartovacia hodnota parametra q_1 na získanie jeho odhadu podmienenou ML metódou bola zvolená na úrovni rozptylu odchýlok, ktoré boli vypočítané aplikáciou metódy najmenších štvorcov na časový rad ročných priemerov úspor obyvateľstva s odhadmi parametrov z lineárneho trendu.

Aplikácia Kalmanových rekurzii na hodnoty výstupu $\{Y_{v+1}, Y_{v+2}, \dots, Y_{v+N}\}$ so špecifikovanými maticami F, G, Q, R poskytne odhady

$$\left\{ \hat{X}_{v+3|v+2}, P_{v+3|v+2}, \hat{X}_{v+2}, P_{v+2|v+2} \right\}, \left\{ \hat{X}_{v+2|v+1}, P_{v+2|v+1}, \hat{X}_{v+1|v+1}, P_{v+1|v+1} \right\}, \dots, \\ \left\{ \hat{X}_{N+1|N}, P_{N+1|N}, \hat{X}_{N|N}, P_{N|N} \right\}$$

Hodnoty odhadov stavových premenných $\hat{X}_{t,1}$ a $\hat{X}_{t,2}$ pre $t = 5, 6, \dots, 24$ sú graficky znázornené na obrázkoch 2 a 3.

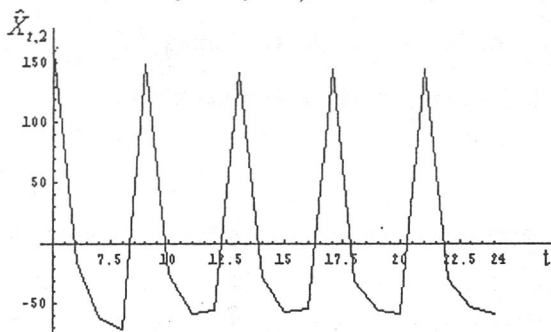
Obrázok 2

Grafické zobrazenie odhadu trendu $\hat{X}_{t,1}$ Kalmanovým filtrom



Obrázok 3

Grafické zobrazenie odhadu sezónnej zložky $\hat{X}_{t,2}$ Kalmanovým filtrom



Analogicky aplikáciou rekurzívnych procedúr Kalmanovho vyrovnávania získame vyrovnané hodnoty stavových premenných $\hat{X}_{t,1}$ a $\hat{X}_{t,2}$. Ich grafický priebeh je zobrazený na obrázkoch 4 a 5.

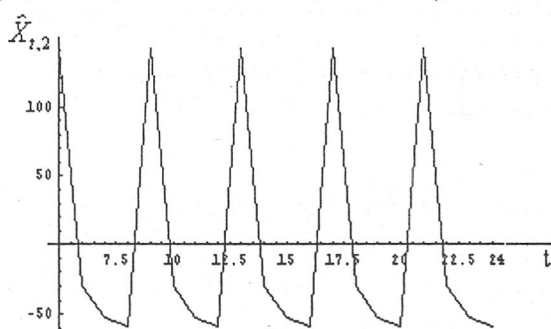
Obrázok 4

Grafické zobrazenie vyrovnaných hodnôt trendu $\hat{X}_{t,1}$ – Kalmanove vyrovnávanie



Obrázok 5

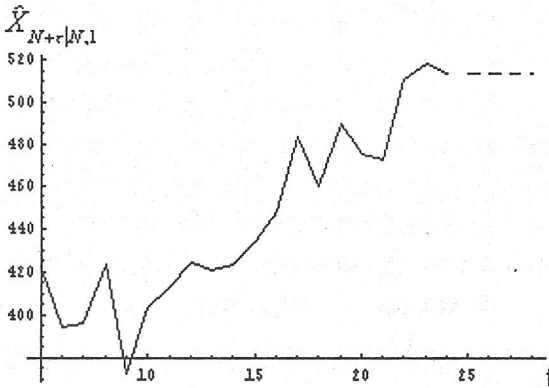
Grafické zobrazenie vyrovnaných hodnôt sezónnej zložky $\hat{X}_{t,2}$ – Kalmanove vyrovnávanie



Nakoniec sa aplikáciou procedúr Kalmanovej predikcie pre horizont $\tau = 4$ získajú hodnoty odhadov stavových premenných $\{\hat{X}_{N+1|N}, P_{N+1|N}\}, \dots, \{\hat{X}_{N+\tau|N}, P_{N+\tau|N}\}$. Predikované hodnoty trendu $\{\hat{X}_{N+\tau|N,1}\}$ a sezónnej zložky $\{\hat{X}_{N+\tau|N,2}\}$ sú graficky znázornené na obrázkoch 6 a 7.

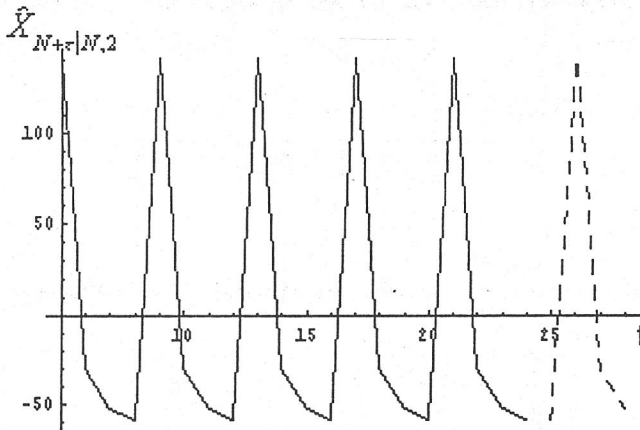
Obrázok 6

Grafické zobrazenie predikovaných hodnôt trendu $X_{N+\tau|N,1}$ – Kalmanova predikcia



Obrázok 7

Grafické zobrazenie predikovaných hodnôt sezónnej zložky $X_{N+\tau|N,2}$ – Kalmanova predikcia



Odhady počiatkových hodnôt stavového vektora \hat{X}_{10} kovariančnej matice P_{10} , odhad rozptylu podmienenou metódou ML, realizáciu Kalmanových rekurzií a grafické zobrazenie vývoja odhadnutých stavových premenných sme vykonal pomocou softvéru *Mathematica* [6].

5. Zhodnotenie a záver

Zaujímavá je presnosť vyrovnania Kalmanovými rekurziami. Vypočítaný index determinácie Kalmanovými rekurziami zo stavovej reprezentácie modelu je $R^2 = 0.99998$, čo je o málo lepší výsledok (prakticky sa približuje k funkčnej závislosti), ako poskytuje klasický štatistický model podľa [2] v tvare

$$y_t = D_t + S_t + \varepsilon_t \quad \text{pre } t = 1, 2, \dots, 24 \quad (24)$$

podľa ktorého index determinácie $R^2 = 0.998328$. V modeli (24) D_t vyjadruje deterministickú časť schémy, S_t je sezónny alebo oscilačný komponent schémy a ε_t je chybový, poruchový alebo náhodný komponent schémy.

Sezónny komponent S_t bol zvolený za predpokladu, že jeho vplyv v každom jednotlivom štvrťroku vo všetkých rokoch má rovnaký účinok. V dôsledku toho, napríklad v j -tom štvrťroku každého roka, jeho hodnoty pre jednotlivé štvrťroky sú b_j , $j = 1, 2, 3, 4$. Možno ho teda interpretovať ako sezónny výkyv charakteristický pre j -tú sezónu. Hodnoty výkyvov modelovanej veličiny v jednotlivých sezónach b_j , $j = 1, 2, 3, 4$ možno taktiež považovať za parametre stojace pri umelých premenných $x_t = (x_{t1} \ x_{t2} \ x_{t3} \ x_{t4})$, $t = 1, 2, \dots, 24$ s hodnotami

$$x_{ij} = 1 \text{ pre } t = j + 4k$$

kde

$$j = 1, 2, 3, 4$$

$$k = 0, 1, 2, \dots, 5$$

$$x_{ij} = 0 \text{ inak}$$

Na záver k modelom stavových premenných ukážeme pre tento príklad vytvorenie štruktúrneho modelu lokálne lineárneho trendu, ale s proporcionálnou sezónnosťou pre časový rad úspor obyvateľstva $\{y_t\}$ [4]. Model je daný rovnicami (7) až (9).

Kalmanovými rekurziami získame rovnaké hodnoty odhadov stavových premenných (trendu a sezónnej zložky) ako v modeli s konštantnou sezónnosťou. Je to preto, lebo výrazy pre Kalmanove rekurzie sú závislé len od počiatkových hodnôt X_t , P_t daných hodnotami výstupov y_1, y_2, \dots, y_t a len od matíc F , G , Q , R . Tie sú v oboch prípadoch rovnaké. Kalmanovými filtráciami, na základe týchto počiatkových podmienok, sa v každom iteračnom kroku vypočítajú nové hodnoty parametrov týchto matíc, t. j. ide o odhad časovo premenlivých (dynamických) parametrov, vzhľadom na získavanie odhadov hodnôt stavových premenných s minimálnou strednou štvorcovou chybou.

Vcelku možno použitie stavových foriem modelov s riešením Kalmanovými rekurziami pre modelovanie a predikciu vývoja úspor obyvateľstva hodnotiť ako veľmi dobré. Dôkaz o tom poskytuje miera determinácie a jej porovnanie

s hodnotou alternatívneho modelu, v našom prípade aditívneho modelu s deterministickou a sezónnou zložkou. Takéto porovnávacie kritérium je veľmi prísne. Dobré výsledky aplikácie stavovej formy modelu a jeho riešenia Kalmanovými rekurziami sa získali aj pri modelovaní inflácie v Českej republike pre časové rady s mesačnou periodicitou [5].

Došlo 4. 5. 2001

Literatúra

- [1] KALMAN, R. E.: A New Approach to Linear Filtering and Prediction Problems. Trans. ASME, J. Basic Eng., 83D, 1960, s. 35–45.
- [2] KOZÁK, J. – SEGER, J.: Jednoduché statistické metódy v prognostice. Praha, SNTL 1971.
- [3] MARČEK, D.: Niektoré prístupy k modelovaniu chaotických ekonomických procesov. Ekonomické rozhlady, 2000, č. 2, s.135–143.
- [4] MARČEK, D.: Analýza, modelovanie, prognózovanie časových radov s aplikáciami v ekonomike. Žilina, EDIS (v tlači).
- [5] VAŠÍČEK, O.: Dynamic Inflation Model. The Analysis and the Forecast of the Macroeconomic Desinflation Strategy. [Bulletin.] The Czech Econometric Society, 1999, č. 10, s. 43–49.
- [6] Wolfram Research, Inc.: Time Series Pack, Reference an User's Guide. Applications Library. Mathematica, marec 1995.

SEASONALITY MODELLING OF TIME AND SAVINGS DEPOSITS OF HOUSEHOLDS BY STATE-SPACE MODELS

Dušan MARČEK

Lots of econometric models discussed in the literature are the models that contain trend and error components only. We now present the method seldom used for modelling trend (deterministic component), seasonal and error component using the dynamic models in state-space form. Model parameters are time varying Their development is estimated adaptively by the Kalman recursions. In recent years quantitative systems based on the state-space representation have been used.

Based on the fact that economic and financial time series of quarterly and monthly or high frequency data contain seasonality, we do not take the seasonality into account, because the econometric model and forecasting method become only second best. A widely used method to incorporate the seasonality into the econometric models is the approach known as the decomposition method.

Unfortunately, the decomposition method has some disadvantage. The decomposition method is basically intuitive and there are a number of theoretical weaknesses in this approach and in estimating model parameters. State-space representation and the associated Kalman recursions have had a profound impact on time series modelling and many related areas.

In this paper we investigate the application of state-space models and the Kalman recursions to quarterly time and savings deposits of households forecasting and stress of the problems encountered with this modelling technique. The time and savings deposits of households is a model with a quarterly time step. Original data are quarterly time series providing a total of 24 observations (see [2]).

For this data, we can express the underlying model in the version with the local linear trend and constant seasonality by the following model equations system

$$\begin{aligned} y_t &= x_{t,1} + x_{t,2} \\ x_{t,1} &= x_{t-1,1} + v_t \\ x_{t,2} &= \sum_{j=1}^{s-1} x_{t-j,2} \end{aligned}$$

where v_t is the white noise variable with zero mean and σ_v^2 variance. The first equation is a transition equation. By third equation is modeled seasonal component suggesting that the seasonal component is the same any one season each year. The first two equations are also the behavioral equations, the next equations are identities.

The above equations system can be represented in linear state-space form, i. e., the discrete series $\{y_t\}$ satisfies the equations of the form

$$\begin{aligned} Y_t &= G_t X_t + W_t \\ X_t &= F_t X_{t-1} + V_t \end{aligned}$$

where X_t is the $\nu \times 1$ vector of inputs (state) or explanatory variable (in our case $X_t [x_{t,1}, x_{t,2}, x_{t-1,2}, x_{t-2,2}]$), ν is number of inputs, Y_t is the $w \times 1$ vector of observed (measured) or explained variables (in our case $Y_t = \{y_t\}$), w is the number of outputs. W_t , V_t are the $\nu \times 1$ or $w \times 1$ vectors of independent Gaussian white noise variables with zero mean and covariance matrices by $E(V_t V_t') = Q_t$ and $E(W_t W_t') = R_t$ respectively.

The first state space equation defines a sequence of observations. The second one is interpreted as describing the evolution of the state X_t of a system at time t in terms of a known sequence of $\nu \times \nu$ matrix F_t . F , G are known matrices or vector at time t . Assuming that the coefficients of the matrices F , G and the initial estimates of the unobservable state vector X are known, one may to obtain an estimate of the state vector $X_{t|t}$ based on the information of Y_t available at time t , i. e., based on the information of Y up to Y_t . Denote the current period by s . We wish to estimate the state vector X at period t based on the information at time s . Let $\hat{X}_{t|s}$ represents the estimate of X_t based on the information

up to and including period s and $P_{t|s}$ represents its mean square error matrix. Kalman recursions find out the best linear estimate of the state vector X_t based on the observations Y_1, Y_2, \dots, Y_s . Then the appropriate set of Kalman recursions for the estimate $X_{t|s}$ in terms of:

$s = t$ is defined as filtering recursion,

$s > t$ is defined as the smoothing recursion, and

$s < t$ is defined as the prediction recursion.

As mentioned earlier the first v values of data were used to calculate the starting values of X and P . To estimate the variance of the random component v_t in the behavioral equations, the log likelihood function of the state-space representation was used in the following form

$$\ln L(\theta; Y_1, Y_2, \dots, Y_N) = -\frac{1}{2} \sum_{t=1}^N \ln \left| \sum_{t|t-1} \right| - \frac{1}{2} (Y_t - \hat{Y}_{t|t-1})' \sum_{t|t-1}^{-1} (Y_t - \hat{Y}_{t|t-1}) + \text{const}$$

where $\sum_{t|t-1}^{-1}$ is the inverse of the error matrix $\sum_{t|t-1} = E \left[(Y_t - \hat{Y}_{t|t-1})(Y_t - \hat{Y}_{t|t-1})' \right]$

Now assuming that the seasonal component of the state variable $x_{t,2}$ in equation (1) is multiplicative – the magnitude of the seasonal swing is proportional (a ratio) to the trend $x_{t,1}$. We can define the version of the structural model with multiplicative seasonal movements over time with respect to the local linear trend and constant seasonality model by

$$y_t = x_{t,1} + x_{t,2} / x_{t,1}$$

$$x_{t,1} = x_{t-1,1} + v_t$$

$$x_{t,2} = \sum_{j=1}^{s-1} x_{t-j,2} / x_{t-j,1}$$

Similarly these equations can be represented in linear state-space form with

$$\mathbf{X}_t = \left[x_{t,1}, x_{t,2} / x_{t,1}, x_{t-1,2} / x_{t-1,1}, x_{t-2,2} / x_{t-2,1} \right]'$$

$$\mathbf{X}_{t-1} = \left[x_{t-1,1}, x_{t-1,2} / x_{t-1,1}, x_{t-2,2} / x_{t-2,1}, x_{t-3,2} / x_{t-3,1} \right]'$$

The Kalman recursions give the some estimates of the state values (trend and seasonal component) as in the case of the model with additive seasonal movements over time. This is because the trend $x_{t,1}$, seasonal component $x_{t,2}$ and the estimations of $X_{t,1}, X_{t,2}$ given y_1, y_2, \dots, y_N is given by the Kalman filter. These estimations obtaining by Kalman filtering are dependent only on the initial values of X_t, P_t and the known matrices F, Q, G, R only and in this case both the information set y_1, y_2, \dots, y_N and the matrices F, Q, G, R are the same.

For comparison purposes, we have generated forecast and computed the coefficient of the determination R^2 for the time and savings deposits of households using the decomposition method, assuming an additive model. The values of R^2 of the state-space models and the Kalman recursions are slight better.