

ARIMA — TRUMPALAIKIO PROGNOZAVIMO PRIEMONĖ

J. MARKELEVIČIUS

Tobulinant statistinės analizės metodologiją esminio statistikos pertvarkymo laikotarpiu, nelieka nuošalyje nauji statistiniai matematiniai metodai, jų taikymas. Viena iš svarbesnių statistinės metodologijos vystymosi sferų yra trumpalaikis prognozavimas, kurio praktiškumas, ypač spartaus socialinio-ekonominio vystymosi sąlygomis, apčiuopiamesnis negu vidutinių, tuo labiau ilgalaikių prognozių. Tai paaiškinama kur kas didesniu trumpalaikių svyravimų, palyginti su kitais, pastovumu ir svarbumu juos reguliuoti arba prie jų prisitaikyti. Tam reikalingas išankstinis jų pažinimas ir modeliavimas.

Be klasikinių socialinių-ekonominų reiškinų dinamikos modelių, pagrįstų trijų žinomų dinamikos eilutės komponentų ryšiu (adityviu, multiplikatyviu arba mišriu), pastaruoju metu užsienio statistikoje plačiai taikoma vadinamoji ARIMA (Autoregressive Integrated Moving Average) metodika. Nušviečiant metodikos koncepciją, galima pasakyti, kad jos taikymas apima du esminius etapus (kaip ir tradicinėje statistikoje): modelio parengimą ir jo panaudojimą prognozuojant. ARIMA koncepcija pirmame etape numato daugiapakopį modelių įvertinimo ir išrinkimo procesą.

Pirmame darbo etape reikia atlikti vadinamąjį dinamikos eilutės „filtravimą“, t. y. žinomų dinamikos dėsningumų (trendo ir sezoninių svyravimų) eliminavimą. Vadinasi, ekonominio ar socialinio reiškinio dinamikos eilutė, kurios lygiai nepasikartoja („originalūs“), kaip yra biologijoje, astronomijoje, eksperimentinėje fizikoje ir t. t., pertvarkoma į stacionaraus proceso pavidalą.

Paprasčiausias filtravimo sprendimas yra pradinės dinamikos eilutės pertvarkymas į grandinių jos rodiklių eilutę, pavyzdžiui:

$$U_t = \Delta Y_t = Y_t - Y_{t-1} \quad (\text{trendo eliminavimas}),$$

$$U_t = \Delta_s Y_t = Y_t - Y_{t-s} \quad (\text{periodinių svyravimų eliminavimas}).$$

Taip transformuota eilutė atvaizduoja trumpalaikių veiksnių įtaką, kuri turi atsispindėti modelyje. Si transformuota eilutė turi tenkinti nors silpno stacionarumo sąlygas (nėra trendo, variacijos lygis laike nekinta, nėra aiškių periodinių svyravimų).

Stacionarumo nustatymo paprasčiausi būdai yra šie: grafinis, autokoreliacijos funkcijos nagrinėjimas, specialūs testai. Griežtai stacionariais vadinami procesai, kurie yra atsitiktiniai, kartais vadinami „baltoju triukšmu“ (asocijuojasi su baltos šviesos spektro esme).

Išfiltruotas procesas (U_t) aprašomas su vadinamųjų ARMA (Autoregressive Moving Average) modelių pagalba. Tai mišraus autoregresinio modelio ir slenkančio vidurkio modelio tipas. Šios klasės modelių struktūrą sudaro trys dalys: autoregresinė, slenkančių vidurkių ir paklaidų. Tokie modeliai vadinami autoregresiniais; jų bendras atvejis gali būti užrašytas taip:

$$\hat{U}_t = \Phi_1 U_{t-1} + \Phi_2 U_{t-2} + \dots + \Phi_p U_{t-p} + E_t.$$

Jie yra p eilės autoregresinis modelis ir žymimi AR (p). Tokio tipo modeliai su daugybe parametru buvo taikomi ir anksčiau. Šioje koncepcijoje $p \leq 2$.

Baigtinio q eilės slenkančio vidurkio modelio MA (q) konkreti reikšmė yra suprantama kaip nepriklausomų q skaičiaus atsitiktinių dydžių tiesinė kombinacija:

$$U_t = E_t + Q_1 E_{t-1} + Q_2 E_{t-2} + \dots + Q_q E_{t-q}.$$

Svorių Q_1, \dots, Q_q suma nebūtinai lygi vienetui. Paprastai kaip ir p , taip ir $q \leq 2$.

Norint pagerinti modelio adaptacines savybes, abu procesai sujungiami ir gaunamas $(p+q+2)$ parametru mišrus ARMA modelis:

$$\Phi(B)U_t = Q(B)E_t.$$

Iš ARMA modelio integracijos būdu gaunamas ARIMA prognostinis modelis, t. y. grįžtama nuo transformuotos dinamikos eilutės (u_t) prie pradinės (y_t).

Ypatinga ARIMA koncepcijoje tai, kad orientuojamasi į trumpalaikius procesus, galimas didelis variantų skaičius.

ARIMA algoritmas apima identifikaciją, įvertinimą, aprobavimą ir retrospektyvinį prognozavimą.

Prognozavimo praktikoje ARIMA metodiką galima derinti su dekompoziciniiais dinamikos eilučių modeliais.

Šią metodiką rekomenduojama taikyti tais atvejais, kai trendo ir sezoninių svyravimų išskyrimas nėra kompliktuotas, o daugiausia dėmesio sutelkiama trumpalaikiams dėsningumams aprašyti. Taigi visų pirma svarbu įvertinti šių dėsningumų įtaką tiriamam procesui. Metodika taip pat numato tiek vienmatį, tiek ir daugiamatį modeliavimą.

Statistiškai procesas pradedamas vertinti nuo didžiausių ARMA modelio eilių, panaudojant Akaikės kriterijų, kuris leidžia nustatyti kompromisą tarp aproksimacijos tikslumo ir modelio parametru skaičiaus. Praktiškiausi yra AR ir MA pirmos ir antros eilių procesai.

ARIMA metodika apima tokius etapus:

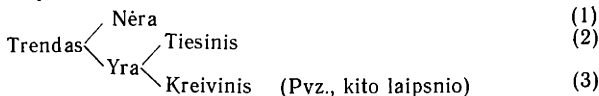
- 1) pirminį dinamikos eilučių tyrimą panaudojant įvairius testus. Pagrindinis tikslas — patikrinti stacionarumą;
- 2) dinamikos eilutės filtravimo procesą (gali būti taikomi įvairių eilių (d) baigtiniai skirtumai);
- 3) atpažinimo funkcijų skaičiavimą. Tai sudaro $U_t = \Delta^d y_t$ proceso identifikacijos pagrindą;
- 4) parametru įvertinimą;
- 5) modelio adekvatiškumo patikrinimą. Tuo tikslu atliekamas likutinių dydžių tyrimas;
- 6) statistinį prognozavimą.

Konceptualiai apžvelgsime šį algoritimą.

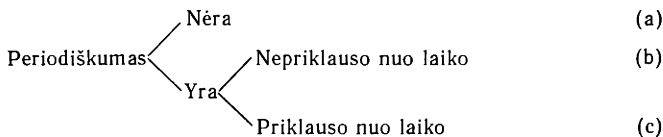
Dinamikos eilutė pradedama tirti panaudojant jos grafinio atvaizdavimo ir kitų žinomų dinamikos dėsningumų atpažinimo priemones. Ypač didelę reikšmę turi grafikų metodas tiriant mažai pažįstamus procesus. Kreivė leidžia nustatyti ne tik trendo buvimą, bet ir periodinius svyravimus. Toliau skaičiuojami dinamikos eilutės rodikliai. Jie padeda nustatyti eilutės stacionarumo sąlygas. Stacionarumui nustatyti dažnai naudojamos dinamikos eilutės lygių kvadratų ir gretimų lygių sandaugos. Šis testas jautrus pasiskirstymo normališkumui.

Kitas testas, pasiūlytas Kocho ir Stiuarto, grindžiamas pirmos ir trečios eilutės dalių vidutinių lygių sugretinimu.

Pirmame etape tiriant eilutės trendą, gali būti trys atvejai:



Tiriant periodiškumo komponentą —



Priklausomai nuo abiejų komponentų kombinacijos formos galima skirti įvairias skurtumų rūšis:

	1	2	3
a	Δ^0	Δ^2	Δ^{K+1}
b	Δ_s	$\Delta\Delta_s$	$\Delta^K\Delta_s$
c	Δ_s^2	Δ_s^2	$\Delta^{K-1}\Delta_s^2$;

čia Δ^K — skirtumo operatorius ($\Delta y_t = y_t - y_{t-1} = (1-B)y_t$).

Jeigu jo nėra, tai dinamikos eilutės pirmiam apdorojimui dažnai gerai tinka pirmos eilės baigtiniai skirtumai.

Dažnai ekonominėms dinamikos eilutėms taikomas $\Delta\Delta_s$ atvejis. Anksčiau eilių skirtumus taikyti gana rizikinga, kadangi tai gali sukelti vadinamąjį „perfiltravimą“, kuris gerokai sumažina eilutės narių skaičių ir blogina prognozės rezultatus. Geriausiais filtrais laikomi tie, kurie korelogramą pakeičia taip, kad tik retais atvejais lieka esminės autokoreliacijos funkcijos reikšmės.

Trečiame etape atliekama modelio identifikacija panaudojant atpažinimo funkcijas. Pagal jų pobūdį galima daryti išvadą apie stochastinio proceso tipą dinamikos eilutėje. Nereikia užmiršti, kad apskaičiuotos ir teorinės atpažinimo funkcijos nevisiškai sutampa dėl „triukšmų“ įtakos. Sie nesutapimai tuo didesni, kuo trumpesnė dinamikos eilutė ir didesni „triukšmai“.

Tai pats sudėtingiausias etapas, kadangi reikia žinoti tiriamo proceso ekonominį turinį, turėti tyrimo metodikos taikymo patirties. Čia reikia subjektyviai nuspręsti, kokį ARMA modelį atitinka gautos atpažinimo funkcijos.

Procesui atpažinti gali būti taikomos šios funkcijos: autokoreliacijos (AKF), dalinės autokoreliacijos (DAKF), porinės autokoreliacijos (PAKF), spektrinio tankio (STF). Svarbiausios yra pirmosios dvi. Jos padeda atpažinti gryno AR arba MA procesų eiles. Kadangi empirinės ir teorinės atpažinimo funkcijos gali iš esmės skirtis, tai, norint patikrinti vizualiai, reikia, be abejo, didelės patirties, kad galima būtų identifikuoti gryną ARMA modelį. Abejotiniais atvejais tikslinga taikyti paprasčiausius standartinius modelius, o sprendimas priimamas juos visus išbandžius. Pavyzdžiui, galima sudaryti AR modelį su nedaugeliu parametru, o likutiniams dydžiams taikomas MA modelis, testuojantis autokoreliacinę seką.

Porinė autokoreliacijos funkcija (PAKF) gali pakeisti dalinės autokoreliacijos funkciją. Tačiau dėl jos nevienareikšmiškumo, susijusio su laiko lango dydžio variacija, praktikoje ji mažiau taikoma.

Svarbią informaciją modeliui identifikuoti duoda spektrinio tankio funkcija. Ji nusako dominuojančius periodinius svyravimus. Jų periodą

arba svyravimo dažnį nurodo spektrinio tankio funkcijos maksimumai (2).

Jeigu pastebimas funkcijos pikas su nuliniu dažnumu, vadinasi, trendas dar nevisiškai eliminuotas, t. y. pažeistas stacionarumas. Apskritai gautas spektras nurodo, kokį filtrą reikia panaudoti transformuotos dinamikos eilutės stacionarumui garantuoti.

Pritaikius pasirinktą filtrą, iš anksto įvertinami „startiniai“ modelio parametrai. Kai jie gerai atitinka atpažinimo funkcijas, tai šiam tikslui galima pritaikyti netgi jų parametrus. Konkretūs parametrai įvertinami 2σ variacijos zonoje.

Pažymėtina, kad modelio parametrai nėra nepriklausomi. Specialiais testais ši priklausomybė eliminuojama išjungiant iš modelio vieną iš jų.

Tiriant likutinius dydžius, svarbiausias reikalavimas yra jų atsitiiktinumas. Tačiau praktiškai neegzistuoja nė vienas sprendimas, garantuojantis vienareikšmį atsakymą. Sprendimas priimamas visumos testų pagrindu. Visų pirma nagrinėjami esminiai nukrypimai nuo nulio. Praktiškai esminiais galima traktuoti nukrypimus, didesnius negu 2σ . Be to, tikrinama jų autokoreliacija ir pasiskirstymo normališkumas. Jų pagrindu ir priimamas sprendimas dėl modelio tinkamumo.

Tikrinant modelį, lieka neišspręstas klausimas apie jo parametru skaičių. Labai svarbu suderinti dvi savybes: parametru skaičiaus saikingumą ir pakankamą modelio išraiškingumą. Šią problemą galima spręsti nuosekliai įtraukiant vis daugiau parametru. Pakankamu parametru skaičiumi galima laikyti tą, kai papildomas naujų parametru įtraukimas likutinių dydžių variacijos iš esmės nesumažina. Be to, tikslinga papildomai sugretinti skirtingos struktūros modelių kokybinius parametrus.

ARIMA metodikos paskirtis gauti prognostinius įvertinimus. Ji leidžia gauti trumpalaikius ir vidutinius įvertinimus. Kai kuriais atvejais, kai prognozių horizontas atitinka dinamikos eilutės ilgį, galima gauti stabilius įvertinimus. Žinoma, jeigu tokioje eilutėje nėra trendo trūkių. Priešingu atveju prognozės horizontas turi būti mažesnis.

Vieno žingsnio prognozės atveju nustatomos prognozuojamo proceso taškinės ir intervalinės reikšmės. Prognozės kokybiniu matu gali būti imamas „pataikymų“ koeficientas, kuris išreiškiamas faktinių reikšmių, nepatenkančių į nustatytą intervalą, skaičiaus santykiu su bendru prognozuojamų reikšmių skaičiumi, kuris nustatomas retrospektyvinės prognozės (jau žinomų dinamikos eilutės narių) pagrindu. Taip gali būti pritaikytas Teilo atitikimo koeficientas.

Norint taikyti ARIMA metodiką, reikia žinoti pagrindines jos sąlygas:

— skaičiavimo operacijų apimtis didelė, būtina taikyti ESM. Be to, tikslinga dirbti dialogo režimu, kuris gerokai patogesnis sudarant prognozavimo modelius ir tikrinant jų adekvatiškumą;

— ARIMA metodiką tikslinga taikyti dinamikos eilutėms, turinčioms ne mažiau kaip 50 narių. VDR statistikų patirtis rodo, kad ši metodika geriausiai tinka eilutėms, charakterizuojančioms dieninių, savaitinių arba mėnesinių lygių dinamiką (3).

Vėlgi ši patirtis patvirtina tai, kad dažniausiai ARIMA metodika pasižymi mažesne prognozės paklaida palyginti su slenkančių vidurkių, eksponentinio išlyginimo, Giunterio ir Smito metodais.

ARIMA metodika buvo panaudota prognozuojant Lietuvos TSR mažmeninę prekių apyvartą (jos mėnesinių apimčių dinamikos pagrindu). Palyginti su klasikiniais trumpalaikio prognozavimo metodais, ARIMA metodika pasitvirtina tais atvejais, kai proceso dinamiką lemia daug veiksnių (bendra respublikos bei konkrečių prekybos organizacijų mažmeninė prekių apyvarta). Prognozuojant tokių maisto produktų, kaip cukrus, sviestas, nealkoholiniai gėrimai, bei nemaisto prekių, t. y. televizorių, avalynės, pardavimą, sezoniško indeksų, apskaičiuotų remiantis autoriaus metodika, tikslumas buvo didesnis (įvertinant pagal likutinių

dydžių variacijos lygį). Tačiau personalinių kompiuterių spartaus diegimo laikotarpiu ARIMA metodika verta prognozavimo specialistų dėmesio ir turi prasmę tolesnis jos plėtojimas.

LITERATURA

1. Бокс Дж., Дженкинс Г. Анализ временных рядов, прогноз и управление. Вып. 1.— М., 1974.
2. Маркелявичюс Й. Ю. Применение спектрального анализа при исследовании внутривременной динамики социально-экономических явлений // Модели анализа и планирования показателей уровня жизни.— М., 1987.
3. Половников В. А., Горчаков А. А. Методы и модели экономического прогнозирования: Учебное пособие.— М., 1980.
4. Götze W. Methodik zur Analyse und Vorausberechnung kurzfristiger betriebswirtschaftlicher Prozesse. Dargestellt am Beispiel der Lebensmittelindustrie.— Berlin, 1985.
5. Mohr W. Univariate—Autoregressive—Moving-Average—Prozesse und die Anwendung der Box-Jenkins-Technik in der Zeitreihenanalyse.— Würzburg, 1976.
6. Rohland U. Programmbeschreibung PP „Multiplikative ARIMA-Modelle“.— Leipzig, 1985.

Vilniaus universitetas
Statistikos katedra

АРИМА — СРЕДСТВО КРАТКОСРОЧНОГО ПРОГНОЗИРОВАНИЯ

И. МАРКЕЛЯВИЧЮС

Резюме

Развитие методологии статистики требует поиска и развития новых направлений. Одним из нетрадиционных подходов к краткосрочному прогнозированию является АРИСС-методика, представляющая собой синтез авторегрессии и скользящего среднего при построении моделей краткосрочного прогнозирования. Разработка модели представляет собой многоступенчатый процесс выбора и оценивания вариантов. Особенности концепции заключаются в том, что в начале проводится фильтрация рядов динамики, т. е. исключаются известные закономерности и моделированию подвергаются остатки ряда динамики, в которых лежат кратковременные закономерности и случайный компонент.

Важная особенность также — использование опознавательных функций: автокорреляции, частной автокорреляции, парной автокорреляции, спектральной плотности. На их основе выбирается соответствующий вариант модели. Это требует, однако, большого опыта.

Важным свойством АРИСС-моделирования является возможность нахождения компромисса между точностью аппроксимации и числа параметров модели.

Прогностическая модель выводится из АРСС (авторегрессия — скользящая средняя) путем интеграции, т. е. производится возвращение от трансформированного ряда к исходному.

Опыт статистиков ГДР показывает, что АРИСС-методика часто обеспечивает более точные и надежные прогностические оценки относительно традиционных методов.

Стремительное развитие компьютерной техники позволяет широко внедрять методику в практику прогнозирования и развивать ее далее.