PREDIKCE ČASOVÉ ŘADY POMOCÍ AUTOREGRESNÍHO MODELU

Ing. Roman DANEL, Ph.D.

roman.danel@volny.cz

Listopad 2004

1. Časové řady

Data, která vytvářejí časovou řadu, vznikají jako pozorování, uspořádané chronologicky (v čase).
S tím se můžeme setkat jak v technice, společenských vědách nebo v ekonomice.

Cílem analýzy časové řady je většinou konstrukce odpovídajícího modelu.
To umožní porozumět mechanismu, na jehož základě jsou sledované údaje generovány.

Znalost modelu umožňuje předpovídat budoucí vývoj systému. Konstrukce modelu umožňuje řídit a optimalizovat činnost příslušného systému vhodnou volbou vstupních parametrů a počátečních podmínek

Volba metody pro analýzu časových řad závisí především na
1. účelu analýzy
2. typu časové řady
3. zkušenosti statistika

S daty, uspořádanými do časové řady souvisí některé specifické problémy. Jsou to:
1. problémy s volbou časových bodů pozorování (na jedné straně je nežádoucí počet pozorování „zhušťovat“ z hlediska složitosti numerických výpočtů, na druhé straně nesmí být pozorování příliš „řídká“ aby nemohl uniknout některý charakteristický rys řady)
2. problémy s kalendářem (řada je zatížena nežádoucími kalendářními nepravidelnostmi)
3. problémy s délkou časových řad – některé metody analýzy vyžadují určitou minimální délku řady, na druhé straně je nebezpečí, že se průběhem času podstatně mění charakteristiky modelu, takže jeho sestavení se stává obtížnější (je nutné připustit např. změny parametrů modelu v čase atd.)

2. Dekompozice časových řad

Časové řady mohou být rozloženy na několik složek:

1. trend
2. sezónní složka
3. cyklická složka
4. reziduální (náhodná) složka

Trend odráží dlouhodobé změny v průměrném chování časové řady. Vzniká v důsledku působení sil, které systematicky působí ve stejném směru.

Sezónní složka popisuje periodické změny v časové řadě, které se odehrávají během jedné sezóny (např. kalendářního roku) a každou další sezónu se opakují. Sezónní složka je způsobena takovými faktory jako např. střídání ročních dob, v ekonomické oblasti např. lidské zvyky apod..

Cyklická složka je periodická složka, která je někdy popisována jako fluktuace okolo trendu (střídání fází růstu a poklesu). Délka a intenzita cyklů může být proměnlivá. Cyklická složka může být důsledkem vnějších vlivů, vytipování příčin ale může být obtížné. Typickým příkladem cyklické složky je obchodní cyklus v ekonomice (konjunktura – krize – deprese)

Náhodná složka je tvořena náhodnými pohyby (fluktuacemi) v průběhu časové řady, které nemají rozpoznatelný systematický charakter. Pokrývá také chyby měření nebo chyby výpočetní (zaokrouhlování). Z pohledu statistiky se jedná o bílý šum s normálním rozdělením.

Zatímco trend, sezónní a cyklická složka jsou funkcí času, náhodná složka je posloupností náhodných veličin.

Časovou řadu si tedy můžeme představit jako trend, na který jsou „nabaleny“ periodické složky (sezónní a cyklická) a bílý šum.

Pro dekompozici je k dispozici několik přístupů:

Klasické dekompoziční metody kladou důraz na systémové složky a jednotlivá pozorování se obvykle berou jako nekorelovaná. Často používaným nástrojem v dekompozičních metodách je regresní analýza.

Boxova-Jenkinsova metodologie dekompozice časové řady bere naproti tomu za základní prvek konstrukce modelu časové řady reziduální (náhodnou) složku, která může být tvořena korelovanými (závislými) náhodnými veličinami. Těžiště spočívá právě ve vyšetřování těchto závislostí, v tzv. korelační analýze.
Nejjednodušší model, s nímž tato metodologie pracuje, je tzv. model klouzavých součtů – MA. Je vhodný pro časovou řadu, v níž jsou všechna pozorování navzájem nekorelovaná až na bezprostřední sousední dvojice. Jinými typy modelů jsou autoregresní modely AR nebo smíšené modely ARMA. Pro řady se zjevným trendovým nebo sezónním charakterem metodologie používá integrované modely ARIMA nebo tzv. sezónní modely, v nichž sezónní nebo trendová složka může být modelována stochasticky. Boxovy-Jenkinsovy modely jsou daleko flexibilnější než dekompoziční modely (flexibilitou modelu rozumíme schopnost modelu adaptovat se rychle na změněný charakter časové řady).

Zcela odlišný přístup analýzy časových řad je **spektrální analýza** (někdy se označuje také jako fourierovská analýza).
Zkoumanou časovou řadu považuje za (nekonečnou) směs sinusových a kosinusových křivek s různými amplitudami a frekvencemi. Pomocí speciálních statistických nástrojů je možné při spektrální analýze získat obraz o intenzitě zastoupení jednotlivých frekvencí v časové řadě (tj. o spektru řady). Lze vytipovat frekvence, které jsou v dané řadě zastoupeny nejvýrazněji a explicitně odhadnout koeficienty periodických složek odpovídajících těmto frekvencím. Spektrální analýza je vhodná při srovnávání chování několika řad, umožňuje nalézt časové zpoždění mezi řadami.

3. Boxova-Jenkinsova metodologie

Obecně můžeme jakýkoli lineární model zapsat ve tvaru:

$$y = G.u + H.e \qquad (3.1)$$

kde

y	výstupní veličina
u	vstupní veličina
G	přenosová funkce - vyjadřuje dynamické vlastnosti systému, tj. závislost výstupu na vstupu
e	bílý šum (rušení) – vyjadřuje skutečnost, že se stejným vstupem dostaneme různý výstup
H	„noise model“ (jak působí rušení na výstup ze standardizovaného zdroje Rušení e(t))

Při analýze časové řady v rámci Boxovy-Jenkinsovy metodologie je postup možno rozdělit do tří kroků:

1. Identifikace
2. Odhad parametrů modelu
3. Ověřování modelu

Pro vybudování spolehlivého modelu podle této metodologie je třeba mít k dispozici alespoň 50 pozorování [1].
Praktická aplikace metodologie je dosti nákladná a časově náročná. Proto se předpokládá použití počítače.

Metoda předpokládá autokorelační vlastnosti časových řad. Stacionarita časové řady znamená, že chování této řady je v jistém smyslu stochasticky ustálené. Stacionární proces je rovnoměrně vyvážený (tj. s konstantním rozptylem), přičemž závislost mezi dvěma libovolnými pozorováními závisí pouze na jejich vzájemné časové vzdálenosti a nikoli na jejich skutečném časovém umístění v řadě. Chování autokorelační funkce je důležitým ukazatelem, protože napovídá, jaký typ modelu je pro danou časovou řadu vhodné použít – toto chování *identifikuje příslušný model*.

Autoregresní model (dále označovaný jako AR) je speciální případ obecného lineárního modelu podle Box-Jenkinsovy metodologie. V praxi se kromě AR modelu požívají ještě také procesy MA (proces klouzavých součtů) nebo ARMA (smíšený proces). Tyto procesy jsou konstruovány co nejúsporněji, aby obsahovaly čím jak nejmenší počet parametrů. Volba parametrů se omezuje tak, aby byla zaručena podmínka stacionarity a invertibility modelu – blíže viz rozbor AR modelu.

4. AR model

Autoregresní model časové řady je založen na poznatku, že každá hodnota v časové řadě, je v relaci (závislosti) s předchozími hodnotami této řady.

Autoregresní model AR(p) řádu p je definován takto:

$$y_t = b_1 y_{t-1} + b_2 y_{t-2} + \cdots + b_p y_{t-p} + \varepsilon_t \qquad (4.1)$$

kde $b_1, b_2 \cdots b_p$ … koeficienty autoregresního procesu

ε_t … bílý šum (současná hodnota)

y_t … nová hodnota řady vypočtená na základě předchozích hodnot

nebo pomocí operátoru zpětného posunutí B lze AR model zapsat jako

$$b(B).y_t = \varepsilon_t \qquad (4.2)$$

kde b(B) je tzv. autoregresní operátor.

Pozn.
Operátor zpětného posunutí **B** je pro lineární proces definován takto:

$$B.y_t = y_{t-1} \tag{4.3}$$

a lze jej aplikovat vícenásobně jako

$$B^j y_t = y_{t-j} \tag{4.4}$$

z toho pak vyplývá zápis AR(p) procesu jako (4.2).

Proces AR(p) odpovídá lineárnímu procesu zapsanému již v invertovaném tvaru (s konečným počtem nenulových parametrů). Proces AR(p) je stacionární, jestliže všechny koeficienty polynomu b(B) leží vně jednotkového kruhu (v komplexní rovině).

*Pod **stacionaritou časové řady** rozumíme to, že charakter této řady zůstává v čase stejný (včetně udržování konstantní úrovně) [1, str. 123]*

***Invertibilní proces** nazýváme lineární proces, který umožňuje zápis ve tvaru, kdy jeho současná hodnota je vyjádřena pomocí minulých hodnot a současné hodnoty bílého šumu. Invertibilita má stěžejní význam pro konstrukci předpovědí.* [1]

Střední hodnota stacionárního procesu AR(p) je nulová.

Autokorelační funkce AR(p) procesu má tvar

$$\rho_k = \alpha_1 G_1^{-k} + \cdots + \alpha_p G_p^{-k} \qquad k >= 0 \tag{4.5}$$

kde α …. jsou konstanty
G …..jsou kořeny polynomu b(B), které jsou navzájem různé a leží vně jednotkového kruhu

Autokorelační funkce procesu AR(p) je tedy v podstatě lineární kombinací klesajících geometrických posloupností a sinusoid různých frekvencí s geometricky klesajícími amplitudami.

Pro výpočet parametrů procesu AR(p) pomocí hodnot jeho autokorelační funkce se používá tzv. Yuelova-Walkerova soustava lineárních rovnic:

$$\begin{aligned}
\rho_1 &= b_1 + b_2\rho_1 + \cdots + b_p\rho_{p-1} \\
\rho_2 &= b_1\rho_1 + b_2 + \cdots + b_p\rho_{p-2} \\
&\cdots \\
\rho_p &= b_1\rho_{p-1} + b_2\rho_{p-2} + \cdots + b_p
\end{aligned} \tag{4.6}$$

Řešením (4.6) dostaneme parametry **b** procesu vyjádřené pomocí hodnot autokorelační funkce.

Zpracování signálů pomocí lineárních modelů je alternativou k zpracování časové řady pomocí spektrální analýzy (rychlé Fourierovy transformace). Rozdíly mezi oběma metodami se snižují zvyšováním řádu AR modelu.

5. Další typy lineárních modelů (MA, ARMA, ARIMA…)

5.1. Model klouzavých součtů MA

Proces MA(q) - z anglického Moving Average - řádu q lze zapsat takto:

$$y_t = \varepsilon_t + w_1\varepsilon_{t-1} + w_2\varepsilon_{t-2} + \cdots + w_q\varepsilon_{t-q} \qquad (5.1)$$

kde w .... jsou parametry modelu
ε_t … bílý šum

nebo pomocí operátoru zpětného posunutí

$$y_t = w(B)\varepsilon_t \qquad (5.2)$$

Proces MA(q) je stacionární pro libovolnou volbu jeho parametrů a jeho střední hodnota je nulová.
Proces MA je invertibilní, jestliže všechny kořeny polynomu w(B) leží vně jednotkové kruhu (v komplexní rovině).
Řešením soustavy nelineárních rovnic autokorelační funkce MA lze získat hodnoty parametrů w.

5.2. Smíšený model ARMA

Kombinací procesů AR(p) a MA(q) dostaneme smíšený proces ARMA(p,q).

Proces ARMA(p,q) můžeme zapsat pomocí symboliky operátoru zpětného posunutí B jako

$$b(B)y_t = w(B)\varepsilon_t \qquad (5.3)$$

Podmínka stacionarity procesu ARMA je shodná s podmínkou stacionarity procesu AR(p) a podmínka invertibility procesu je shodná s podmínkou invertibility procesu MA(q).
Střední hodnota procesu ARMA je nulová a jeho autokorelační funkce vyhovuje podobné soustavě diferenčních rovnic jako v případě AR procesu.

Pro zápis procesů ARMA (nebo AR) lze použít několik alternativních možností
a) ve tvaru diferenční rovnice
b) ve tvaru lineárního procesu
c) ve invertibilním tvaru

Pro každý AR model řádu p lze nalézt ekvivalentní MA model s dostatečným počtem q elementů rušení.

Některé ekonomické nebo obchodní časové řady mohou být modelovány s celkem malým počtem p a q prvků v podobě modelu AR, MA nebo modelu ARMA.

Cílem je najít nejmenší počet p a q prvků, potřebných k uspokojivé předpovědi časové řady.

5.3. Integrovaný smíšený model ARIMA

Nejobecnějším modelem je nestacionární integrovaný smíšený model ARIMA(p,d,q), kde „I“ znamená integraci. První vyhodnocení modelu ARIMA vytvořil J. Durban v roce 1960. ARIMA modely umožňují popis procesů, u kterých nejenže dochází ke změnám úrovně, ale tyto změny mohou mít nesystematický, náhodný charakter (což je pro časové řady z praxe běžné). Tento model modeluje stochasticky vedle náhodných fluktuací i trendovou složku. Modely ARMA se konstruují až pro řadu prvních nebo vyšších diferencí časové řady, u konstrukce ARIMA modelu nepožadujeme stacionaritu analyzované řady.

Pro úplnost uvádím definici modelu ARIMA(p,d,q)

$$b(B).v_t = w(B).\varepsilon_t \qquad (5.4)$$

kde

$$v_t = \Delta^d y_t \qquad (5.5)$$

je d-tá diference modelovaného procesu y pro proces v (kde proces v je vlastně proces ARMA(p,q) s nenulovou střední hodnotu).

Diferenční operátor je možno vyjádřit pomocí operátoru zpětného posunutí jako

$$\Delta = 1 - B \qquad (5.6)$$

6. Výstavba modelů

6.1. Identifikace modelu

Je první fází výstavby modelu. Úkolem je rozhodnout, jaký typ modelu použít (AR, MA, ARMA) a explicitně určit řád modelu.

Jestliže střední hodnota stacionárního procesu není nulová, je nutné danou řadu centrovat. O případné nenulovosti střední hodnoty lze rozhodnout na základě nějakého statistického testu (srovnáme aritmetický průměr, což chápeme jako odhad střední hodnoty s dvojnásobkem odhadnuté směrodatné odchylky).

Vlastní identifikace je založena na zkoumání průběhu odhadnuté autokorelační funkce a parciální autokorelační funkce [1], tabulka str. 124 a125. Snažíme se zjistit existenci případného identifikačního bodu k0 (tj. hodnota, za kterou autokorelační funkce začíná být nulová) nebo zjistit, že taková hodnota k vůbec neexistuje. Chování autokorelační funkce napovídá, jaký typ modelu je vhodné pro danou řadu použít

6.2. Odhad parametrů modelu

V současné době existuje řada odhadových procedur. Jejich popis je značně komplikovaný a k jejich provedení je zapotřebí použít počítače (použijí se hrubé odhady parametrů, získané během identifikace a ty se dále zpracovávají v iteračních procedurách. Základem těchto procedur je jsou odhadové metody nejmenších čtverců).

Některé z možných metod odhadu parametrů autoregresního modelu AR(p) jsou:
a) Least-squares method (metoda nejmenších čtverců)
b) Yule-Walker method
c) Burg's method („maximální entropie") – doporučovaná metoda pro modelování širokopásmových signálů [4]
d) Forward-backward
e) Maximum Likehood Extimator (metoda maximální věrohodnosti)

Příkladem software pro odhad parametrů modelu je programový prostředek Matlab (moduly „Systém Identification Tool", „Time Series Analysis Toolbox", atd.) firmy MathWorks nebo program pro statistické výpočty Systat firmy SPSS.

6.3. Ověřování modelu

Ověřování (diagnostika) modelu má za úkol potvrdit nebo zamítnout adekvátnost modelu. Existuje několik metod ověření modelu, jako jsou např. metoda přeparametrizování modelu, metoda odhadnutých reziduí nebo Portmanteu test, test pomocí kumulativního periodogramu apod. Podrobný popis těchto metod je uveden v publikaci [1], strany 136 až 140. Každá metoda posuzuje model z různých hledisek a s různou účinností, doporučuje se proto použít více metod současně.

Podezření na neadekvátnost modelu je například tehdy, jestliže jsou neúměrně velké standardní odchylky odhadnutých parametrů nebo je neúměrně velká hodnota rozptylu σ_ε^2. V tom případě použijeme model s větším počtem parametrů (vyššího řádu). Jestliže nový model vede ke snížení směrodatné odchylky odhadů a odhady nově zavedených parametrů jsou výrazně nenulové, pak původní model nahradíme novým rozšířeným modelem. Tento postup je principem metody přeparametrizování modelu, uvedené v odstavci výše.

7. Použití AR modelů v praxi

Kde se používá aproximace AR modelem?
- pro parametrickou spektrální analýzu
- fyzikální modely (jako je např. rozpoznávání řeči)
- modely se sinusovým průběhem
- signálové filtry
- modelování vyvíjejících se širokopásmových signálů (signály, jejíž spektra se v čase vyvíjejí – např. únavový test materiálu)
- zpracování signálu z EKG nebo vyšetřování mozkové aktivity

Další informace k použití AR modelu viz [7].

Příkladem praktického použití AR modelu je restaurace vysokofrekvenčních složek v hudebním signálu, které se ztratily např. dlouhodobým skladováním na magnetofonové pásce nebo restaurace poškozeného signálu při nahrávání ze starých gramofonových desek

U modelování signálů pomocí autoregrese je vysoká rozlišovací schopnost metody AR značně závislá na přítomnosti šumu. Když se poměr signál/šum zmenšuje, je zapotřebí k získání stejného rozlišení vyšší řád modelu. Příliš velký řád modelu ale způsobuje chyby ve vypočteném spektru. AR modely nejsou příliš vhodné pro identifikaci jednoduchých, sinusových, spektrálních složek. Pro analýzu přechodových signálů je lépe použít waveletovou transformaci [4].

8. Literatura

1. Cipra, T.: Analýza časových řad s aplikacemi v ekonomii. Praha, SNTL 1986.
2. Enochson L. – Otnes K.: Analiza numeryczna szeregów czasowych. Warszawa, Wydawnictwo Naukowo Techniczne 1978.
3. The System Identification Toolbox – Matlab, dokumentace, 1998.
4. Kuběna R.: Analýza signálů pomocí autoregresního modelování. Sborník XXI. Semináře ASŘ'98, příspěvek č.21.
5. Stanford R.: Time Series Forecasting Models (chapter 12), Econometric Forecasting Techniques (chapter 13) & appendix 3 „ARIMA modeling" – http://www.furman.edu/~stanford/res.htm, Furman University, Greenwille.
6. Tůma J.: Použití autoregresního modelu k aproximaci impulsní charakteristiky. Sborník XXI. Semináře ASŘ'98, příspěvek č.53.
7. Mc Graw F.: Parametric Spectral Estimation. http://www.ms.washington.edu/~s520/chpt-09/, 1998.
8. http://www.mathtools/toolboxes.htm – Time Series Analysis Toolbox, 1999.

Příloha 1 - Význam parametru FPE při práci s AR modely v software Matlab

FPE (Final Prediction Error) je kritérium, které se používá k určení řádu autoregresního modelu. Jeho autorem je Akaike a patří spolu s AIC kritériem (Akaike Information Criterion) mezi nejstarší a nejčastěji používaná kritéria.

Kritérium FPE pro stanovení řádu autoregresního modelu má tvar:

$$FPE(k) = \log \sigma_k^2 + (2k)/n$$

kde n je délka analyzované řady

Pro rostoucí parametr k - AR(k) od nuly po správný řád p modelu AR(p) má odhad rozptylu klesající tendenci, po překročení p kolísá kolem správné hodnoty rozptylu. Založit konstrukci odhadu řádu modelu přímo na minimalizaci rozptylu by vedlo k neúměrnému preferování velkých hodnot k. Proto se používá penalizační funkce w(k), která penalizuje volbu příliš vysokého řádu modelu.

Odhad řádu modelu obecně získáme minimalizací odhadového kritéria. Řád modelu se určuje na základě kompromisu mezi malými hodnotami penalizační funkce a velkými hodnotami odhadu rozptylu.

Řád autoregresního modelu se na základě kritéria FPE určuje tak, aby při něm byla minimální chyba předpovědi o jeden krok dopředu (pracuje se v něm s tzv. asymptotickou střední čtvercovou chybou předpovědi). Tato předpověď je založena na odhadnutém modelu právě posuzovaného řádu a na daných pozorováních $y_k, k = 1, \cdots, n$.

Formálně je FPE speciální případ kritéria AIC.