PROGNOSTICKÉ METODY P2

2009-02-26

Vlastnosti časových řad, modely časových řad:

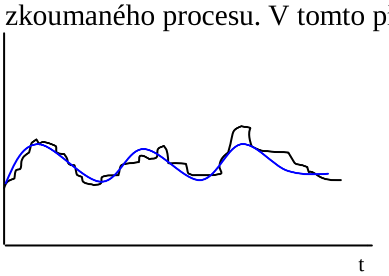
- ✓ **Časová řada** – věcně a prostorově srovnatelná pozorování (dat), která jsou jednoznačně uspořádána z hlediska času ve směru minulost – přítomnost.
- ✓ **Analýza časových řad** – soubor metod, které slouží k popisu těchto řad a případně k předpovídání jejich budoucího chování.

Elementární charakteristiky časových řad:

- ✓ Obvykle prvním úkolem při analýze časové řady je získat rychlou a orientační představu o charakteru procesu, který tato řada reprezentuje.
- ✓ **Vizuální analýza:**
  - Grafy
  - Elementární charakteristiky – diference různého řádu, tempa a průměrná tempa růstu, průměr časové řady aj.

Přístupy k modelování časových řad:

- ✓ Výchozím principem je jednorozměrný model: $Y_t=f(t;u)$, k němuž se v zásadě přistupuje trojím způsobem
- ✓ **Klasický (formální) model:**
  - **Dekompozice časové řady na 4 složky:**
    - $Y=T+S+C+u$
    - $Y_t=T_t+S_t+Ct+u_t$
    - $Y_t=T_t\times S_t\times C_t\times u_t$
  - První tři jsou systematické části časové řady, poslední je nesystematická (náhodná, reziduální)
  - Časová řada nemusí mít všechny tři složky
  - **T – trendová složka**, dlouhodobá složka, vypovídá o dlouhodobé funkci
    - Deterministická
    - Stochastická
  - **C – cyklická složka**, kolísání časové řady, charakter je delší než jeden rok, někteří autoři jí nazývají také střednědobý trend, občas je zahrnuta v trendové složce, fáze expanze, vrchol, recese, dno
  - **S – sezónní složka**, představuje to kolísání časové řady, ale je kratší než jeden rok, čtvrtletní nebo měsíční data, jak s ní zacházet – zda s ní modelovat nebo jí odstranit
    - <u>Deterministická sezónnost</u> – odstraňujeme pomocí dummy proměnných
    - <u>Stochastická sezónnost</u> – odstraňujeme sezónním diferenciálem nebo softwarovým očištěním (metoda CENSUZ), sezónní indexy k očištění časové řady
  - u – náhodná složka: $E(u)=0$, $Var(u)=\sigma^2<\infty$, $Cov(u_t,u_s)=0$, $t\neq s$ => regresní analýza
- ✓ **Box-Jenkinsonova metodologie:**
  - Tento přístup považuje za základní prvek časové řady náhodnou složku a snaží se ji modelovat. Těžiště postupu se klade na korelační analýzu více či méně závislých pozorování, uspořádaných do tvaru časové řady.
  - Náhodná složka ut má charakter bílého šumu, takže platí, že $Cov(u_{t-h},u_t)=0$ pro všechna h
- ✓ **Spektrální analýza:**
  - V tomto přístupu se časová řada považuje za „směs" sinusovek a kosinusovek o rozličných amplitudách a frekvencích. Tato koncepce umožňuje provést explicitní popis periodického chování časové řady a především vystopovat ty významné složky periodicity, které se podílejí na věcných vlastnostech zkoumaného procesu. V tomto přístupu tedy není stěžejním faktorem časová proměnná, ale právě faktor.



DGP (Proces generování dat)
VAR → VECM
↓
Kointegrační analýza

Vlastnosti stochastických časových řad:

- ✓ Stochastický proces lze označit jako nekonečnou posloupnost náhodných veličin uspořádaných v čase
- ✓ Možnost realizace každého pozorování je dána funkcí rozdělení pravděpodobnosti $f(Y_t)$
- ✓ Modelování konkrétního stochastického procesu proto vyžaduje co nejpřesněji popsat povahu náhodnosti:
  - Skutečná povaha zpravidla neznámá.
  - Aproximativně pomocí zjednodušeného modelu časové řady.
- ✓ Čím přesněji popisuje stochastický model časové řady charakteristiky skutečného rozdělení pravděpodobnosti, tím lepší je jeho schopnost predikce
- ✓ **Náhodný proces** – množina náhodných veličin ze stejného pravděpodobnostního prostoru (Ω, J, P) indexovaných t z množiny T (TCR), který reprezentuje čas.

Stacionarita a nestacionarita časových řad:

- ✓ **Veličiny** diskrétní (naše, v určitý okamžik), spojité
- ✓ **Popis stochastického procesu** – společné (simultánní) rozdělení pravděpodobnosti hodnot náhodné veličiny Y_t, tj. $(Y_{R+1}, Y_{R+2}, \ldots, Y_{R+T})$
- ✓ **Stacionarita:**
  - **Definice** – časová řada yt je stacionární, jestliže její rozdělení pravděpodobnosti je v čase neměnné, tj. společné (rozdělení) pravděpodobnosti $(Y_{R+1}, Y_{R+2}, \ldots, Y_{R+T})$ není závislé na R.
    Striktní stacionarita.
  - **Slabě stacionární proces** (2. řádu) – průměr a rozptyl je konstantní, přičemž kovariance libovolných dvou pozorování časové řady závisejí pouze na velikosti zpoždění, tj. na délce časového posunu mezi nimi, nikoliv na hodnotě R.
- ✓ **Nestacionární časové řady:**
  - Trend, sezónnost, strukturální šoky v ekonomice
  - **Zjišťování nestacionarity:**
    - ACF – autokorelací funkce:

      Koeficient korelace k-tého řádu: $\rho_k = \dfrac{\gamma_k}{\gamma_0}$

      $k = 0, \pm1, \pm2, \ldots, \pm k$, $-1<\rho_k<1$

      - *Skutečná:*
        $\gamma_k = \operatorname{cov}(Y_T, Y_{T+k}) = \operatorname{cov}(Y_{T-K}, Y_T) = \gamma_{-k}$
        $\gamma_0 = \sigma_{Y_T} \times \sigma_{Y_{T+k}} = \sigma^2_{Y_T}$
      - *Odhadovaná:*
        K = 1, 2, …, k
        $\hat{\gamma}_k = \sum_{t=1}^{T-k} \left(Y_t - \overline{Y}\right)\left(Y_{t+k} - \overline{Y}\right)$
        $\hat{\gamma}_0 = \sum_{t=1}^{T} \left(Y_t - \overline{Y}\right)^2$
        $\rho_k = 0$ pro všechna k => stacionární
        - Vlastnosti odhadu $\hat{\rho}_k$:
          T>50 a K<T/4, potom odhad $\hat{\rho}_k$ asymptoticky nestranný => můžeme testovat $H_0 = \hat{\rho}_k = 0$ pro k=1,2,..., $\dfrac{T}{4} - 1 = K$

          Společná: $\dfrac{T}{4} - 1 = K$

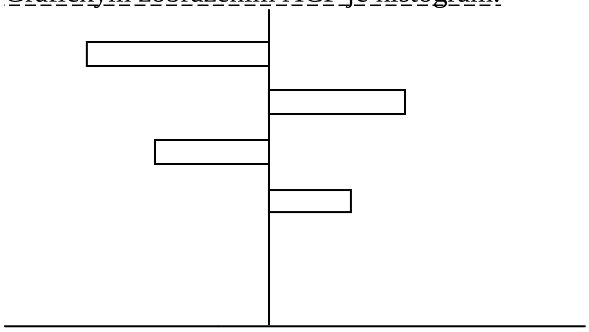
          $H_0 : \hat{\rho}_1 = \hat{\rho}_2 = .. = \hat{\rho}_K$

Testové kritérium => Q: $Q = T\sum_{K=1}^{K}\hat{\rho}_K^2$

χ^2 rozdělení $\hat{\rho}_k = 0$ s K stupni volnosti

Q > Q* nevede na H_0

- Grafickým zobrazením ACF je histogram:



ρ_k

- Stacionární – jestliže korelační koeficient s rostoucím K konverguje k nule
- Nestacionární – jestliže opak

○ Testy jednotkového kořene

MODELY ČASOVÝCH ŘAD:

ARIMA: (p, d, q)

✓ **Stacionární:**
- MA – model klouzavých průměrů, paměť v délce rovné zpoždění
- AR – autoregresní proces, nekonečná paměť
- ARMA – smíšený proces – je-li časová řada agregovaná (stačí agregace s chybou měření)

✓ **Nestacionární:**
- IMA
- ARI
- ARIMA

✓ Speciálně: SARIMNA

ARIMA modely:

✓ Použití ke **krátkodobé** predikci a to v situaci, kdy nejsou k dispozici adekvátní data vysvětlujících proměnných, resp. Když při odhadu a verifikaci LRM nebo MSR dospějeme k závěru, že odhadnuté parametry jsou z hlediska ekonomických statistických i ekonometrických kritérií nepoužitelné. Nebo má-li model špatné prognostické vlastnosti.

Modely náhodných procházek:

✓ Jednoduchý stochastický proces

✓ Proces je nestacionární, neboť průměr a rozptyl nejsou konstantní

✓ Model typu I(1)

Definice:

I(d) – časová řada je integrována řádu d, přičemž d značí, kolikrát se musí časová řada diferencovat, abychom dostali stacionární časovou řadu =>
- I(0) – časová řada je stacionární ve svých původních hodnotách, není nutné diferencovat
- I(1) – časová řada, jejíž první diference jsou stacionární
- Apod.

✓ Použije-li se proces tohoto typu k popisu dynamického chování, např. cen akcií nebo spotřebitelských cen, jde o nestacionární model časových řad.

✓ Avšak časové řady cenových změn jsou již generovány stacionárním ryze náhodným procesem, nazývaným také jako bílý šum.

✓ Pro krátkodobou předpověď na základě modelu náhodné procházky platí:

✓ $Y_t = Y_{t-1} + u_t$

- u_t… u – $E(u_t)=0$, $Var(u_t)=\sigma^2<\infty$, $Cov(u_t,u_s)=0$, $t\neq s$
  $Y_t - Y_{t-1}=u_t$
  $\Delta Y_t=u_t$ =>I(1)
- $\beta=1$… model náhodných procházek
  $\beta<1$… AR model
  AR(1)