Základy ekonometrie

IX. Analýza jednorozměrných časových řad

Obsah tématu

Úvod

- Specifika práce s časovými řadami a problémy:
  1. Vliv časově zpožděných proměnných.
  2. Regrese s *nestacionárními* proměnnými $\rightarrow$ *zdánlivé regrese (spurious regression)*.
- Analýza vlastností časových řad.
- Autoregresní modely $\rightarrow$ test jednotkového kořene (test stochastického trendu).

Základní značení

- Y_t pro $t = 1, \ldots, T$.
- Zpožděné proměnné o p období: Y_{t-p} pro $t = p + 1, \ldots, T$.
- Význam řada vs. pozorování vychází z kontextu.

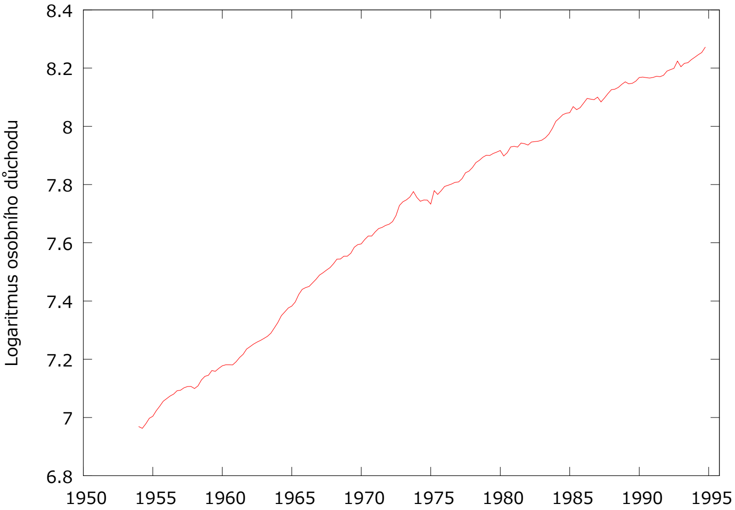
Číslo řádku	Y	Zpožděné Y	Zpožděné Y o dvě období	Zpožděné Y o tři období
1	Y_4	Y_3	Y_2	Y_1
2	Y_5	Y_4	Y_3	Y_2
3	Y_6	Y_5	Y_4	Y_3
4	Y_7	Y_6	Y_5	Y_4
5	Y_8	Y_7	Y_6	Y_5
6	Y_9	Y_8	Y_7	Y_6
7	Y_{10}	Y_9	Y_8	Y_7

Obsah tématu

Úvod

- Nestacionární řady v regresi $\rightarrow$ zdánlivá regrese.
- Stacionarita $\times$ nestacionarita $\rightarrow$ trend.

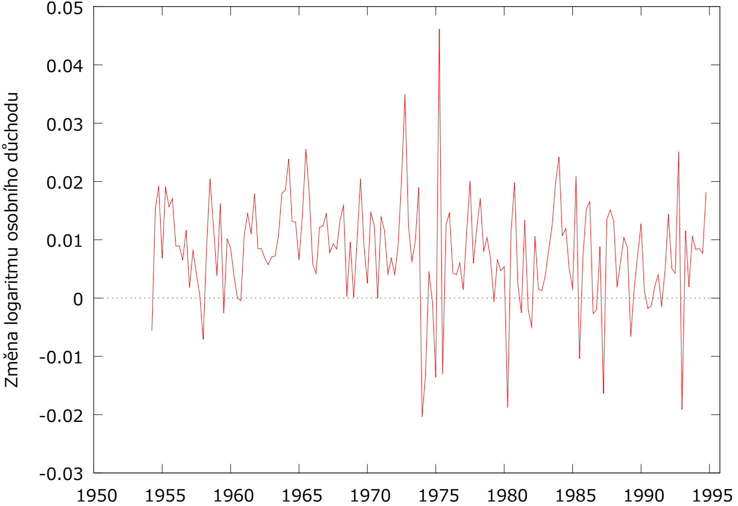
Graf časové řady logaritmu osobního důchodu v USA.



Diferencování

- Obvyklý způsob stacionarizace řady.
- První diference Y_t: $\Delta Y_t = Y_t - Y_{t-1}$.
- Řada v logaritmu $\rightarrow$ diference = tempo růstu.

Graf časové řady změny logaritmu osobního důchodu v USA



Korelační vlastnosti časové řady

- Korelace mezi pozorováními.
- Korelace Y_t a Y_{t-1} (pro osobní důchod) $= 0.999716$.
- Korelace ΔY_t a ΔY_{t-1} (pro osobní důchod) $= -0.00235$.
- Nestacionarita a korelace v čase $\times$ stacionarita a nízká korelace v čase.
- Důležité pro modelování časových řad.

Obsah tématu

Úvod

- Autokorelační funkce: autokorelace jako funkce délky zpoždění (p).

$$r_p = corr(Y_t, Y_{t-p}).$$

- Prakticky: „vyhození" prvních p pozorování (teoreticky není třeba u všech autokorelačních funkcí).

Autokorelační funkce

Délka zpoždění, p	Osobní důchod, Y	Změna v osobním důchodu, ΔY
1	0.9997	-0.0100
2	0.9993	0.0121
3	0.9990	0.1341
4	0.9986	0.0082
5	0.9983	-0.1562
6	0.9980	0.0611
7	0.9978	-0.0350
8	0.9975	-0.0655
9	0.9974	0.0745
10	0.9972	0.1488
11	0.9969	0.0330
12	0.9966	0.0363

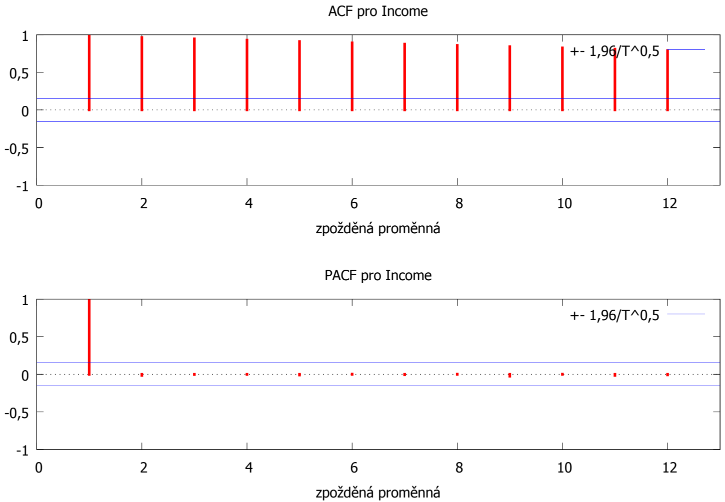
Význam autokorelace

- Y je vysoce korelované v čase $\times$ ΔY tuto vlastnost nemá.
- Znalost minulého důchodu $\rightarrow$ dobrý odhad důchodu dnes $\times$ znalost jen temp růstu.
- Y: chování s *dlouhou pamětí (long memory)*.
- Y nestacionární $\times$ ΔY stacionární.

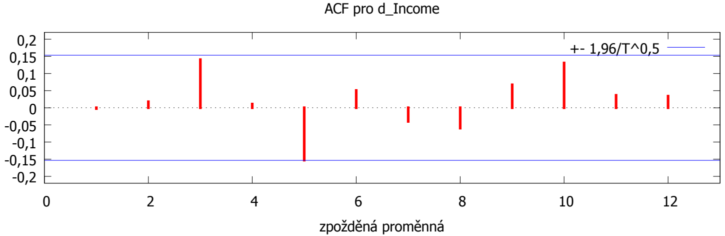
Parciální autokorelační funkce

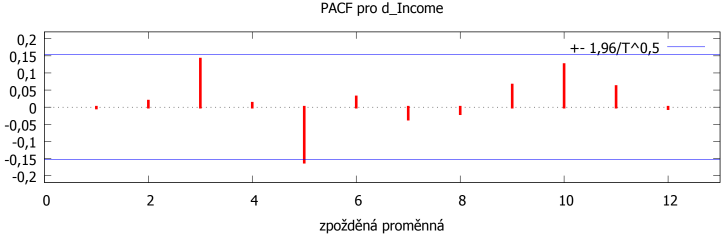
- Parciální autokorelační funkce: korelace jako funkce délky zpoždění (p) při zohlednění (očištění) korelací nižších řádů.
- Prakticky: koeficient z regrese nezpožděné proměnné na zpožděné proměnné do požadovaného řádu zpoždění a uchování koeficientu pro poslední zpožděnou proměnnou (to je příslušný parciální autokorelační koeficient).

ACF a PACF logaritmu osobního důchodu v USA



ACF a PACF změny logaritmu osobního důchodu v USA





Poznámka k použití korelogramu v Gretlu

- Výpočty ACF se liší oproti korelačním koeficientům pro časové řady nezpožděných a zpožděných proměnných.
- Při korelacích: odlišné průměry a směrodatné odchylky.
- Ve výpočtu ACF: použity stejné průměry a směrodatné odchylky pro nezpožděnou řadu.

Obsah tématu

Úvod

- Autokorelační funkce = užitečný nástroj pro souhrn vlastností řady.
- Korelace mají svá omezení → preferovaná regrese.
- Analýza závislosti řady na svých zpožděných hodnotách → autoregresní (AR) modely.

Obsah tématu

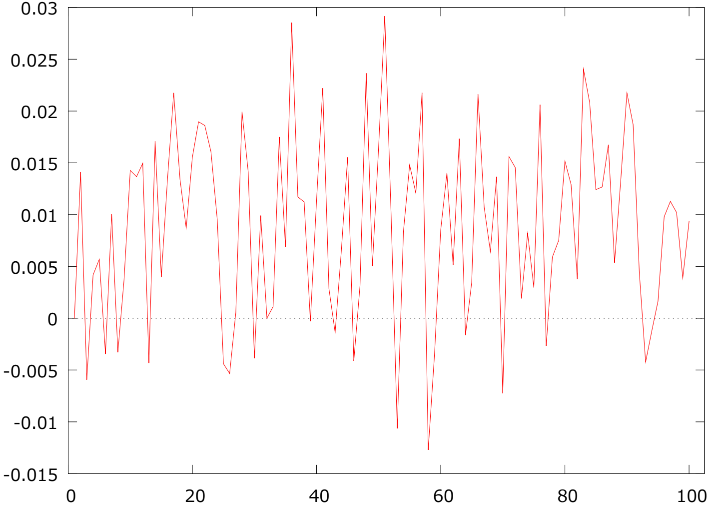
Úvod

- Z analýzy autokorelovaných náhodných složek.
- $AR(1)$ model:

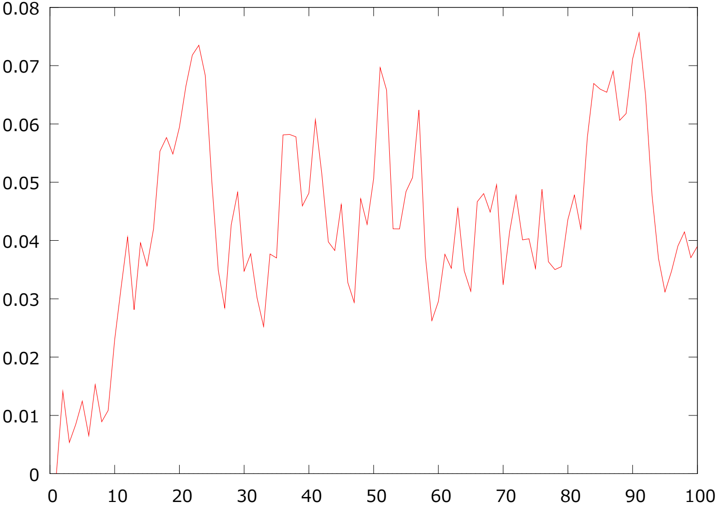
$$Y_t = \alpha + \rho Y_{t-1} + \epsilon_t.$$

- ρ determinuje autokorelační vlastnosti řady (viz problém autokorelace náhodných složek = stejná odvození).
- Ilustrace: $\rho = 0$, $\rho = 0.8$ a $\rho = 1$ (podchycení různých typů chování).
- Y je stacionární, pokud $|\rho| < 1$; nestacionární pro $\rho = 1$ (*jednotkový kořen*) a $|\rho| > 1$ (nestacionární explozivní vývoj).

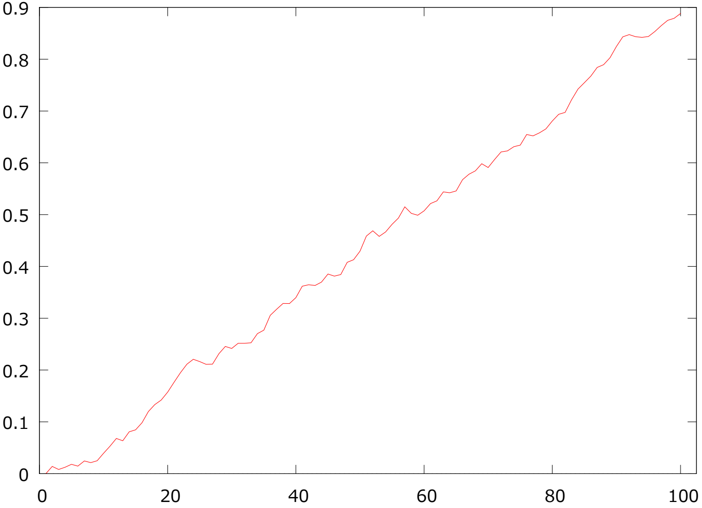
$AR(1)$ časová řada s $\rho = 0$



$AR(1)$ časová řada s $\rho = 0.8$



$AR(1)$ časová řada s $\rho = 1$



$AR(1)$ časová řada s $\rho = 1$

- V $AR(1)$ modelu: $\rho = 1 \rightarrow Y$ nestacionární; $|\rho| < 1 \rightarrow Y$ stacionární.
- Y s jednotkovým kořenem: autokorelace blízko jedné a příliš neklesají.
- Y s jednotkovým kořenem: dlouhá paměť.
- Y s jednotkovým kořenem: trendové chování (hlavně pro nenulové α).
- Y s jednotkovým kořenem: ΔY stacionární $\rightarrow$ *diferenčně stacionární*.

Diferencování $AR(1)$ s jednotkovým kořenem

- Odečtění Y_{t-1} z obou stran rovnice:

$$Y_t - Y_{t-1} = \alpha + \rho Y_{t-1} - Y_{t-1} + \epsilon_t.$$

- Jiný zápis:

$$\Delta Y_t = \alpha + \phi Y_{t-1} + \epsilon_t.$$

- $\phi = \rho - 1$ (pro $\rho = 1$ je $\phi = 0$).
- Test $\phi = 0 \Rightarrow$ test jednotkového kořene časové řady.
- Stacionarity pro $-1 < \rho < 1 \Rightarrow -2 < \phi < 0$ *podmínka stacionarity*.

Model náhodné procházky

- Pro $\rho = 1$ ($\phi = 0$):
$$Y_t = \alpha + Y_{t-1} + \epsilon_t.$$
- Náhodná procházka (random walk) pro $\alpha = 0$ $\times$ náhodná procházka s driftem (random walk with drift) pro $\alpha \neq 0$.

Obsah tématu

Úvod

- Autoregresní model řádu p (pro $t = p+1, \ldots, T$):

$$Y_t = \alpha + \rho_1 Y_{t-1} + \ldots + \rho_p Y_{t-p} + \epsilon_t.$$

- Obvyklé odečtení Y_{t-1}:

$$\Delta Y_t = \alpha + \phi Y_{t-1} + \gamma_1 \Delta Y_{t-1} + \ldots + \gamma_{p-1} \Delta Y_{t-p+1} + \epsilon_t.$$

- Koeficienty sklonu, $\phi, \gamma_1, \ldots, \gamma_{p-1}$, jsou funkce $\rho_1, \ldots, \rho_p$ z původního $AR(p)$ modelu.
- $\phi = 0 \rightarrow AR(p)$ model Y obsahuje jednotkový kořen.
- $-2 < \phi < 0 \rightarrow$ stacionarita.
- Obvyklá stacionarizace řady s jednotkovým kořenem na základě prvních diferencí.

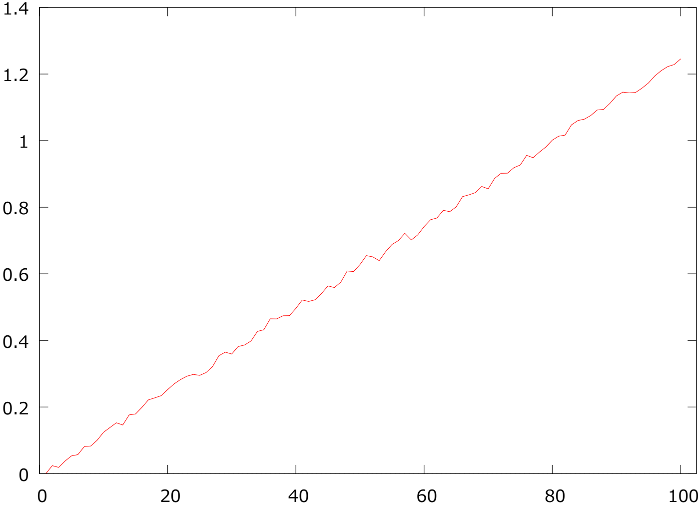
Trendy

- V regresních modelech nechceme řady s trendy (kromě případu tzv. kointegrace).
- Deterministický trend:

$$Y_t = \alpha + \delta t + \epsilon_t.$$

- Stochastický trend = řada s jednotkovým kořenem.
- Kombinace: $Y_t = \alpha + \rho Y_{t-1} + \delta t + \epsilon_t$.

Trendově stacionární $AR(1)$ časová řada ($\alpha = 0$, $\rho = 0.2$)



Shrnutí

- Nestacionární časové řady obsahující jednotkový kořen (stochastický trend); *diferenčně stacionární*.
- Stacionární časové řady s $-2 < \phi < 0$ v rámci $AR(p)$ modelu; možnost trendového chování (deterministický trend) $\rightarrow$ *trendově stacionární*.

$AR(p)$ model s deterministickým trendem

- Obecný model:

$$\Delta Y_t = \alpha + \phi Y_{t-1} + \gamma_1 \Delta Y_{t-1} + \ldots + \gamma_{p-1} \Delta Y_{t-p+1} + \delta t + \epsilon_t.$$

- Proč ne původní specifikace?
  1. Snadné testování jednotkového kořene (parametr $\phi = 0$).
  2. Problém multikolinearity $Y_{t-1}, \ldots, Y_{t-p}$.
- Testování v rámci modelu.

Příklad – osobní důchod v USA

Tabulka: $AR(4)$ model s deterministickým trendem.

Proměnná	OLS odhad	t-statistika	p-hodnota
Konstanta	0.138	1.279	0.203
Y_{t-1}	-0.018	-1.190	0.236
ΔY_{t-1}	-0.017	-0.217	0.829
ΔY_{t-2}	0.014	0.172	0.863
ΔY_{t-3}	0.130	1.627	0.106
t	0.0001	0.955	0.341

Obsah tématu

Testy zahrnující $\alpha, \gamma_1, \ldots, \gamma_{p-1}$ a δ

- Určení optimální délky zpoždění.

Krok 1. Volba maximální délky zpoždění, *pmax*.

Krok 2. Odhad (s využitím OLS) modelu $AR(pmax)$ s deterministickým trendem; test $\gamma_{pmax-1} = 0$.

Krok 3. Odhad $AR(pmax-1)$ modelu. Test významnosti parametru u nejvyššího zpoždění.

Krok 4. Opakovaný odhad AR modelů nižšího řádu $\rightarrow AR(p)$ model, pro který je γ_{p-1} statisticky významné (nebo „dojdou“ zpoždění).

Krok 5. Test zanedbání deterministického trendu: $H_0 : \delta = 0$.

Příklad – osobní důchod v USA

Tabulka: $AR(1)$ model.

Proměnná	OLS odhad	t-statistika	p-hodnota
Konstanta	0.039	2.682	0.008
Y_{t-1}	-0.004	-2.130	0.035

Výběr modelu vs. průměrování modelů

- Kritika postupu sekvenčního výběru.
- Možnost zamítnutí „lepšího modelu“.
- Prezentace výsledků jen jediného modelu?
- Průměrování modelů (hlavně bayesovský přístup); různé váhy dle kvality vyrovnání dat.

Informační kritéria a výběr modelu

- Použitelné pro jakýkoli model, hlavně model časových řad (i korigovaný koeficient determinace jako informační kritérium).
- Typická podoba:
$$IC(\theta) = 2\ln[L(\theta)] - g(p).$$
- p je počet koeficientů modelu a $g(p)$ je rostoucí funkce p.
- θ: odhady OLS nebo ML.
- Hodnota věrohodnostní funkce podobný vývoj jako R^2 (z hlediska přidávání dalších proměnných).

Typická informační kritéria

- *Bayesovské informační kritérium (BIC)* resp. *Schwarzovo kritérium*:

$$BIC(\theta) = 2\ln[L(\theta)] - p\ln(T).$$

- *Akaikeho informační kritérium (AIC)*:

$$AIC(\theta) = 2\ln[L(\theta) - 2p].$$

- *Hannanovo-Quinnovo kritérium (HQ)*:

$$HQ(\theta) = 2\ln[L(\theta)] - pc_{HQ}\ln[\ln(N)],$$

- c_{HQ} je konstanta (doporučené různé volby); HQ je konzistentní kritérium volby modelu pro $c_{HQ} > 2$ (s pravděpodobností jedna vybere korektní model pro nekonečně velký vzorek).

Informační kritéria – alternativní vyjádření

- Pozor při interpretaci, někdy IC jako:

$$IC(\theta) = -2 \ln[L(\theta)] + g(p).$$

- Analogicky BIC, AIC a HQ
- V tomto případě ideální model s minimální hodnotou IC!
- Viz dokumentace ekonometrického programu nebo porovnání hodnot (a znamének) logaritmu věrohodnostní funkce a hodnoty kritérií (znamének).

Testy zahrnující ϕ: testy jednotkového kořene

- Test $\phi = 0 \rightarrow t$-statistika $\times$ nekorektní použití tabulek t-rozdělení (statistika nemá t-rozdělení).
- Intuitivní vysvětlení:

$$\widehat{\beta} \sim N\left(\beta, \frac{\sigma^2}{TE(X^2)}\right).$$

- AR model (náhodná vysvětlující proměnná):

$$\widehat{\phi} \sim N\left(\phi, \frac{\sigma^2}{TE(Y_{t-1}^2)}\right).$$

- $E(Y_{t-1}^2) = var(Y_{t-1}) + [E(Y_{t-1})]^2 \rightarrow$ rozptyl proměnné s jednotkovým kořenem je nekonečno.

Dickeyho-Fullerův test

- t-statistika + využití Dickeyho-Fullerova rozdělení (není standardní rozdělení, hodnoty tabelovány).
- Kritické hodnoty závisí na přítomnosti nebo nepřítomnosti deterministického trendu.
- „Dickeyho-Fullerův test" pro $\phi = 0$ v $AR(1)$ modelu.
- „Augmented Dickeyho-Fullerův test" pro $\phi = 0$ v $AR(p)$ modelu.
- Prakticky:

$$\Delta Y_t = \alpha + \phi Y_{t-1} + \gamma_1 \Delta Y_{t-1} + \ldots + \gamma_{p-1} \Delta Y_{t-p+1} + \epsilon_t.$$

- Volba optimálního řádu zpoždění a rozhodnutí o zahrnutí trendu $\rightarrow$ optimální model + t-statistika $\rightarrow$ porovnání s kritickými hodnotami.

Kritické hodnoty DF testu

Tabulka: Kritické hodnoty Dickeyho-Fullerova testu.

	$T = 25$	$T = 50$	$T = 100$	$t = \infty$
AR model bez deterministického trendu				
1% kritická hodnota	-3.75	-3.59	-3.50	-3.42
5% kritická hodnota	-2.99	-2.93	-2.90	-2.80
AR model s deterministickým trendem				
1% kritická hodnota	-4.38	-4.15	-4.04	-3.96
5% kritická hodnota	-3.60	-3.50	-3.45	-3.41

Problémy a další testy

- „Nízká síla testu“: nalezení jednotkového kořene, i když neexistuje (strukturální zlomy, náhlé výkyvy).
- Implementován v ekonometrických programech.
- Další testy: Phillipsův-Perronův test (stejná nulová hypotéza), KPSS testy (nulová hypotéza o stacionaritě řady kolem deterministického trendu).
- KPSS: Denis Kwiatkowski, Peter C.B. Phillips, Peter Schmidt and Yongcheol Shin (1992).
- Příklad s osobním důchodem v USA.

Obsah tématu

Formální definice:

- Slabá (kovarianční) stacionarita Y:
  1. $E(Y_t)$ je stejné pro všechna t.
  2. $var(Y_t)$ je konečný a stejný pro všechny hodnoty t.
  3. $cov(Y_t, Y_{t-s})$ závisí pouze na s, nikoli na t.
- Silná stacionarity: Y_t má rozdělení konstantní v čase.

$AR(1)$ model

- Model:

$$Y_t = \alpha + \rho Y_{t-1} + \epsilon_t,$$

- Stacionarita pro $|\rho| < 1$:

$$E(Y_t) = \frac{\alpha}{1-\rho},$$

$$var(Y_t) = \frac{\sigma^2}{1-\rho^2},$$

$$cov(Y_t, Y_{t-s}) = \frac{\rho^s \sigma^2}{1-\rho^2},$$

kde $\sigma^2 = var(\epsilon_t)$.

Model s deterministickým trendem

- Model:

$$Y_t = \alpha + \delta t + \epsilon_t.$$

- Lze ukázat:

$$\begin{aligned} E(Y_t) &= \alpha + \delta t, \\ var(Y_t) &= \sigma^2, \\ cov(Y_t, Y_{t-s}) &= 0. \end{aligned}$$

- $E(Y_t)$ závisí na t $\rightarrow$ model nestacionární.
- Stacionární po odstranění trendu $\rightarrow$ trendově stacionární řada.

Obsah tématu

Úvod

- Zájem o volatilitu proměnné.
- Volatilita v modelech z oblasti financí: CAPM, Blackův-Sholesův model oceňování finančních derivátů apod.
- *ARCH* a *GARCH* modely: snadná implementace (náročnější teorie v pozadí).

Obsah tématu

Úvod

- Model náhodné procházky s driftem:

$$\begin{aligned} Y_t &= \alpha + Y_{t-1} + \epsilon_t \\ \Delta Y_t &= \alpha + \epsilon_t \end{aligned}$$

- Model náhodné procházky pro cenu aktiva zcela korektní.
- $\Delta y_t = \Delta Y_t - \overline{\Delta Y}$, kde $\overline{\Delta Y} = \frac{\sum \Delta Y_t}{T}$.

$$\Delta y_t = \epsilon_t.$$

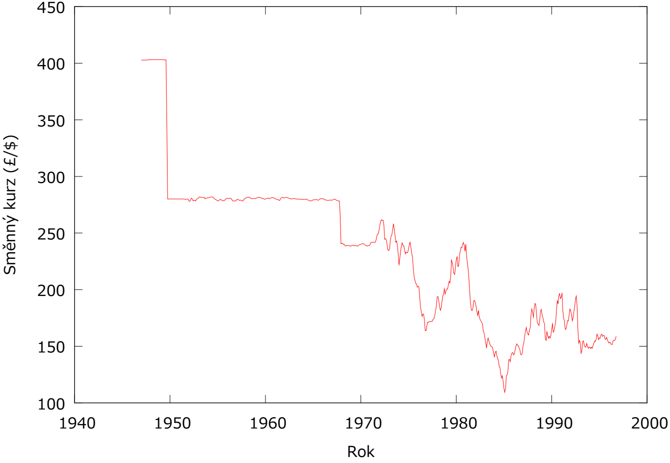
Jednoduché modelování volatility

- Δy_t^2 jako odhad volatility v čase t.
- Měřítko volatility malé ve stabilních časech a velké v časech bouřlivých.
- Alternativní důvod: ze vztahu pro odhad rozptylu ΔY_t s využitím jediného pozorování.
- Volatilita proměnlivá v čase, lze snadno spočítat.
- Modelování shluků ve volatilitě (clustering in volatility): autoregresní modely.
- Např. $AR(1)$ model:

$$\Delta y_t^2 = \alpha + \rho \Delta y_{t-1}^2 + \epsilon_t.$$

- Predikovatelná volatilita, ne samotný výnos! (standardní aplikace technik, rozšíření na $AR(p)$)
- Příklad: vývoj směnného kurzu libry vzhledem k dolaru.

Ilustrace



Obrázek: Časová řada vývoje směnného kurzu GBP vzhledem k USD.

Příklad – cena akcie (1/5)

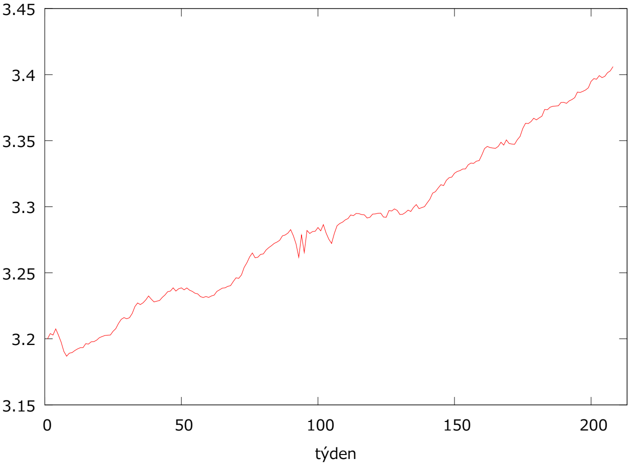
- Vývoj ceny akcie: 208 týdnů (`stock.gdt`).
- Konstrukce Δy_t^2.
- $AR(1)$ model na základě standardního postupu volby optimální délky zpoždění.
- Možnosti predikce vývoje volatility.

Příklad – cena akcie (2/5)

Tabulka: $AR(1)$ model pro Δy_t^2.

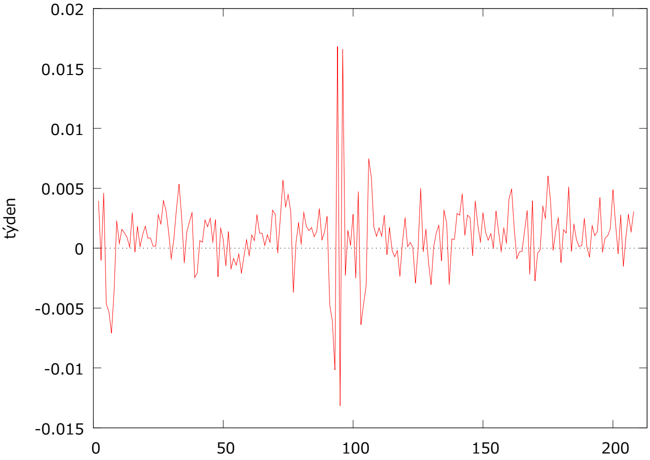
Proměnná	OLS odhad	t-statistika	p-hodnota
Konstanta	0.024	1.624	0.106
Δy_{t-1}^2	0.737	15.552	0.000

Příklad – cena akcie (3/5)



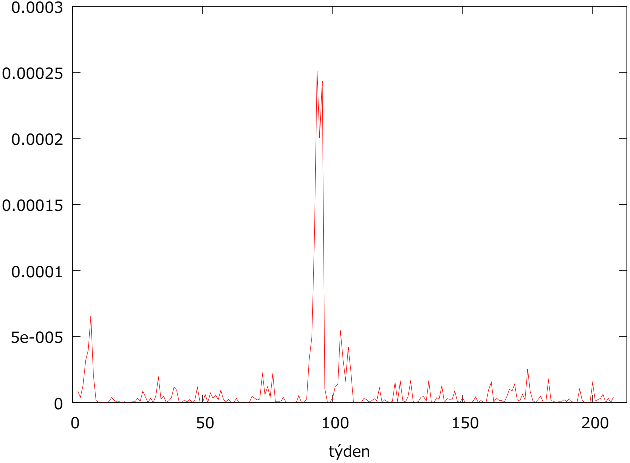
Obrázek: Graf časové řady logaritmu ceny akcie.

Příklad – cena akcie (4/5)



Obrázek: Graf časové řady změny logaritmu ceny akcie.

Příklad – cena akcie (5/5)



Obrázek: Graf časové řady volatility akcie.

Obsah tématu

Úvod

- Obecně (včetně rozšíření) v kontextu

$$Y_t = \alpha + \beta_1 X_{1t} + \ldots + \beta_k X_{kt} + \epsilon_t.$$

- Modely s $\alpha = \beta_1 = \ldots = \beta_k = 0$.
- *ARCH* model se týká rozptylu (volatilitě) náhodných složek, ϵ_t, pro jednoduchost:

$$\sigma_t^2 = var(\epsilon_t).$$

- *ARCH* model s p zpožděními (označovaný jako $ARCH(p)$):

$$\sigma_t^2 = \gamma_0 + \gamma_1 \epsilon_{t-1}^2 + \ldots + \gamma_p \epsilon_{t-p}^2.$$

- Koeficienty odhadovány např. metodou ML.
- Pokud jen centrovaný výnos akcie, Δy_t:

$$\sigma_t^2 = \gamma_0 + \gamma_1 \Delta y_{t-1}^2 + \ldots + \gamma_p \Delta y_{t-p}^2$$

Pokračování příkladu – cena akcie (1/3)

- Možno pracovat přímo s ΔY_t.
- Standardní interpretace.
- Odhady volatility (a predikce).

Pokračování příkladu – cena akcie (2/3)

Tabulka: ARCH(1) model pro data výnosnosti akcií.

Proměnná	Odhad koeficientu	p-hodnota	95% int. spolehlivosti
Regresení rovnice s vysvětlovanou proměnnou ΔY			
Konstanta	0.105	0.000	[0.081;0.129]
ARCH rovnice			
Konstanta	0.024	0.000	[0.016;0.032]
$\Delta\varepsilon_{t-1}^2$	0.660	0.000	[0.302;1.018]

Pokračování příkladu – cena akcie (3/3)

Tabulka: ARCH(2) model pro data výnosnosti akcií.

Proměnná	Odhad koeficientu	p-hodnota	95% int. spolehlivosti
	Regresení rovnice s vysvětlovanou proměnnou ΔY		
Konstanta	0.109	0.000	[0.087;0.131]
	ARCH rovnice		
Konstanta	0.025	0.000	[0.016;0.033]
$\Delta\varepsilon^2_{t-1}$	0.717	0.000	[0.328;1.107]
$\Delta\varepsilon^2_{t-2}$	-0.043	0.487	[-0.165;0.079]

Rozšíření ARCH modelu

- *GARCH*, *SAARCH*, *TARCH*. *AARCH*, *NARCH* nebo *NARCHK*.
- Modely *stochastické volatility* (ne třída *ARCH* modelů).
- *Zobecněný (generalized) ARCH*: *GARCH*.
- Přidání zpožděných hodnot míry volatility, např. $GARCH(p, q)$:

$$\sigma_t^2 = \gamma_0 + \gamma_1 \epsilon_{t-1}^2 + \ldots + \gamma_p \epsilon_{t-p}^2 + \lambda_1 \sigma_{t-1}^2 + \ldots + \lambda_q \sigma_{t-q}^2.$$

- Možnost popsat širší paletu chování časových řad.

Pokračování příkladu – cena akcie (GARCH)

Tabulka: GARCH(1,1) model pro data výnosnosti akcií.

Proměnná	Odhad koeficientu	p-hodnota	95% int. spolehlivosti
	Regresení rovnice s vysvětlovanou proměnnou ΔY		
Konstanta	0.109	0.000	[0.087;0.131]
	GARCH rovnice		
Konstanta	0.026	0.000	[0.015;0.038]
$\Delta\varepsilon^2_{t-1}$	0.714	0.000	[0.327;1.101]
σ^2_{t-1}	-0.063	0.457	[-0.231;0.104]

Obsah tématu

MA modely

- $MA(q)$:

$$Y_t = \epsilon_t - \theta \epsilon_{t-1}.$$

- ϵ_t (inovace) v čase vzájemně nekorelovány (vlastnosti bílého šumu).
- Vlastnosti $MA(1)$:

$$\begin{aligned} E(Y_t) &= 0 \\ var(Y_t) &= \sigma^2(1+\theta^2) \\ cov(Y_t, Y_{t-1}) &= -\theta \\ cov(Y_t, Y_{t-1}) &= 0 \quad \text{pro} \quad s \geq 2. \end{aligned}$$

$MA(q)$ modely

- $MA(q)$:
$$Y_t = \epsilon_t - \theta_1 \epsilon_{t-1} - \ldots - \theta_q \epsilon_{t-q}.$$
- Specifické vlastnosti autokorelační funkce (nulová pro $q+1$).

Vztah *AR* a *MA*

- Stacionární $AR \rightarrow MA(\infty)$.
- Invertibilní $MA \rightarrow AR(\infty)$.
- *Autoregresní model klouzavých součtů*, $ARMA(p,q)$:

$$Y_t = \alpha + \rho_1 Y_{t-1} + \ldots + \rho_p Y_{t-p} + \epsilon_t - \theta_1 \epsilon_{t-1} - \ldots - \theta_q \epsilon_{t-q}.$$

- Kombinace vlastností *AR* a *MA*, velmi flexibilní, možnost deterministického trendu.
- Y_t s jednotkovým kořenem $\rightarrow$ práce s ΔY_t.
- *Autoregresní integrovaný model klouzavých součtů (ARIMA)*:

$$\Delta Y_t = \alpha + \rho_1 \Delta Y_{t-1} + \ldots + \rho_p \Delta Y_{t-p} + \epsilon_t - \theta_1 \epsilon_{t-1} - \ldots - \theta_q \epsilon_{t-q}.$$