SROVNÁNÍ VYBRANÝCH METOD PREDIKCE ZMĚN TRENDU INDEXU PX

Jiří Trešl, VŠE v Praze

1. Úvod

Predikce budoucího vývoje finančních časových řad představuje jeden z nejobtížnějších problémů současné statistiky a ekonometrie (Franses, vanDijk, 2006; Kadem, Fokianos, 2002). „Klasické“ časové řady, se kterými se setkáváme v přírodních a technických vědách, totiž vykazují zpravidla jednodušší průběh a lépe identifikovatelné signály. Totéž platí i pro některé ekonomické časové řady, kdy je možno provést klasickou dekompozici a modelovat trend i sezónní složku. Naproti tomu u finančních časových řad pozorujeme komplikované chování, změny volatility, časté body zvratu a náhlé změny směru trendu. Nepochybně je toto chování dáno superpozicí postojů a investičních strategií individuálních investorů. Zásadní potíž při modelování takových časových řad představují časové změny modelových parametrů a z toho vyplývající neúčinnost predikce prostou extrapolací. Přes nesporný pokrok v oblasti modelování autoregresní podmíněné heteroskedasticity a režimových změn (GARCH, modely s proměnlivými režimy) se zdá, že lepších výsledků bude možno dosáhnout aplikací metod umělé inteligence (McNelis, 2005; Gencay, Stengos, 1998).

Obrázek 1
Časová řada vývoje indexu PX v období 2002–2009

Cílem předložené analýzy je predikce změny trendu v časové řadě indexu PX, reprezentujícího vývoj kapitálového trhu v ČR. Vstupní data tvoří jeho denní ceny v období 2002–2009, tj. 1988 hodnot. V každém obchodním dnu jsou k dispozici

otvírací, uzavírací, minimální a maximální ceny, které budeme značit O (OPEN), C (CLOSE), L (LOW), H (HIGH). Kromě toho budeme používat i denní (O1,C1,L1,H1) a týdenní (O5,C5,L5,H5) logaritmické výnosy v procentech. Např. denní a týdenní logaritmické výnosy vztažené k uzavíracím cenám C byly vypočteny jako

$$C1_t = 100 \ln(C_t / C_{t-1}) \quad C5_t = 100 \ln(C_t / C_{t-5}) \quad (1)$$

Analogicky byly vypočteny rovněž denní a týdenní logaritmické výnosy vztažené k O,L,H.

Z grafu na obrázku 1 je patrné, že hodnoty indexu PX mají rostoucí trend a dosahují prvního maxima v bodě 1366 (C = 1903,80 v čase 09/07/2007). Po krátkodobém propadu je pak absolutní maximum ve sledovaném období dosaženo v bodě 1445 (C = 1936,90 v čase 29/10/2007). Tento bod je možno považovat za zlomový z hlediska nástupu klesajícího trendu. K predikci obou maxim postupně použijeme následující postupy:

- GARCH modely s proměnnými parametry
- Klouzavé charakteristiky
- Cenové indikátory technické analýzy
- Umělé neuronové sítě

2. GARCH modely s proměnnými parametry

Notorickým problémem v modelování finančních časových řad je výskyt autoregresní podmíněné heteroskedasticity, kdy dochází v určitých obdobích k náhlému zvýšení volatility relativních výnosů a poté k její stabilizaci. Dnes již klasickým nástrojem k překonání této obtíže je použití modelů třídy ARCH (Engle, 1982) a zobecněných modelů GARCH (Bollerslev, 1986).

Symetrický GARCH (p,q) lze obecně psát ve tvaru

$$X_t = \sigma_t z_t \quad z_t \approx N(0,1) \quad \sigma_t^2 = \omega + \sum_{i=1}^q \alpha_i X_{t-i}^2 + \sum_{j=1}^p \beta_j \sigma_{t-j}^2 \quad (2)$$

kde σ_t^2 značí podmíněný rozptyl a X_t modelovaný stochastický proces, v našem případě relativní denní výnos. Podmíněný rozptyl tedy závisí na kvadrátech minulých hodnot procesu a na minulých hodnotách podmíněného rozptylu. Další významné zobecnění představuje model EGARCH (p,q), umožňující zachytit případnou asymetrickou reakci na kladné a záporné hodnoty procesu (Nelson, 1991)

$$\ln(\sigma_t^2) = \omega + \sum_{j=1}^p \beta_j \ln(\sigma_{t-j}^2) + \sum_{i=1}^q \left(\alpha_i \left| \frac{X_{t-i}}{\sigma_{t-i}} \right| + \gamma_i \frac{X_{t-i}}{\sigma_{t-i}} \right) \quad (3)$$

V naší analýze se omezíme na základní EGARCH (1,1) model typu

$$\ln(\sigma_t^2) = \omega + \beta \ln(\sigma_{t-1}^2) + \alpha \left| \frac{X_{t-1}}{\sigma_{t-1}} \right| + \gamma \frac{X_{t-1}}{\sigma_{t-1}} \quad (4)$$

Je zřejmé, že pro $\gamma = 0$ mají kladné i záporné hodnoty procesu stejný vliv na volatilitu. Hodnoty $-1 < \gamma < 0$ signalizují, že kladné hodnoty procesu ovlivňují volatilitu méně než hodnoty záporné.

Pro vlastní výpočty byly celé časové řady rozděleny do úseků délky 120 obchodních dnů (tj. 6 obchodních měsíců), v nichž byly odhadnuty parametry modelů EGARCH(1,1). Toto okno bylo použito klouzavým způsobem s posunem 60 obchodních dnů (tj. 3 obchodní měsíce). Celkem tedy vzniklo 32 intervalů 1-120, 61-180, ..., 1801-1920, 1861-1980. Výsledky modelování jsou shrnuty v tabulce 1 a 2. Na obrázku 2 a 3 je znázorněn časový vývoj odhadnutých parametrů.

Tabulka 1
Odhady parametru β EGARCH modelu (4)

Číslo	Interval	C1	p-value	O1	p-value	L1	p-value	H1	p-value
1	1-120	0,846	0,000	0,972	0,000	-0,395	0,641	0,914	0,000
2	61-180	0,948	0,000	0,991	0,000	0,955	0,000	-0,218	0,441
3	121-240	0,951	0,000	0,116	0,818	0,963	0,000	0,029	0,955
4	181-300	0,441	0,255	0,446	0,276	-1,008	0,000	0,645	0,021
5	241-360	0,639	0,094	0,857	0,000	-0,478	0,004	-0,953	0,000
6	301-420	0,761	0,010	0,823	0,000	0,725	0,000	0,807	0,000
7	361-480	0,568	0,509	0,298	0,829	0,976	0,000	-0,168	0,749
8	421-540	0,950	0,000	0,978	0,000	-0,978	0,000	-0,588	0,377
9	481-600	0,921	0,000	0,925	0,000	0,837	0,000	0,854	0,000
10	541-660	0,762	0,000	0,694	0,000	0,892	0,000	0,738	0,000
11	601-720	0,457	0,075	0,386	0,333	0,038	0,918	0,437	0,260
12	661-780	0,605	0,000	0,666	0,000	0,763	0,000	0,730	0,000
13	721-840	0,656	0,000	0,743	0,000	0,747	0,000	0,674	0,000
14	781-900	0,980	0,000	-0,945	0,000	0,973	0,000	0,944	0,000
15	841-960	0,963	0,000	0,896	0,000	0,950	0,000	0,882	0,000
16	901-1020	1,008	0,000	-0,759	0,000	0,949	0,000	0,893	0,000
17	961-1080	1,048	0,000	1,011	0,000	0,952	0,000	-0,775	0,000
18	1021-1140	0,988	0,000	0,990	0,000	0,986	0,000	0,957	0,000
19	1081-1200	0,947	0,000	-0,192	0,308	0,959	0,000	0,950	0,000
20	1141-1260	0,952	0,000	0,897	0,000	-0,621	0,304	0,923	0,000
21	1201-1320	-0,579	0,043	0,849	0,000	-0,409	0,099	0,987	0,000
22	1261-1380	-0,768	0,000	0,907	0,000	0,904	0,000	0,892	0,000
23	1321-1440	0,978	0,000	0,947	0,000	0,918	0,000	0,932	0,000
24	1381-1500	1,011	0,000	-1,024	0,000	0,912	0,000	-0,272	0,764
25	1441-1560	0,961	0,000	0,885	0,000	0,967	0,000	-0,394	0,387
26	1501-1620	0,937	0,000	0,927	0,000	0,954	0,000	0,956	0,000
27	1561-1680	-0,081	0,670	1,084	0,000	0,933	0,000	0,882	0,000
28	1621-1740	0,966	0,000	0,970	0,000	0,978	0,000	0,977	0,000
29	1681-1800	0,953	0,000	-0,694	0,000	0,934	0,000	0,960	0,000
30	1741-1860	0,359	0,356	-0,731	0,061	0,755	0,000	0,811	0,000
31	1801-1920	0,989	0,000	0,997	0,000	1,007	0,000	-0,079	0,950
32	1861-1980	-0,478	0,131	0,287	0,419	0,546	0,057	0,953	0,000

Zdroj: vlastní výpočty v EViews 6

Tabulka 2

Odhady parametru γ EGARCH modelu (4)

Číslo	Interval	C1	p-value	O1	p-value	L1	p-value	H1	p-value
1	1-120	-0,164	0,064	-0,217	0,003	0,121	0,521	-0,262	0,001
2	61-180	-0,112	0,234	-0,085	0,217	-0,122	0,057	0,187	0,234
3	121-240	-0,058	0,568	-0,125	0,442	-0,129	0,201	0,087	0,560
4	181-300	-0,251	0,096	-0,246	0,101	0,176	0,006	0,052	0,696
5	241-360	-0,194	0,227	-0,174	0,024	0,110	0,462	0,098	0,225
6	301-420	-0,097	0,214	-0,102	0,161	-0,181	0,022	-0,007	0,910
7	361-480	-0,009	0,930	-0,078	0,610	0,093	0,007	-0,076	0,611
8	421-540	0,160	0,000	0,069	0,175	0,123	0,000	-0,158	0,239
9	481-600	-0,192	0,008	-0,146	0,000	-0,174	0,081	-0,111	0,174
10	541-660	-0,240	0,091	-0,323	0,005	-0,141	0,080	-0,315	0,003
11	601-720	-0,336	0,010	-0,147	0,333	0,110	0,413	-0,232	0,238
12	661-780	-0,251	0,012	-0,179	0,129	-0,126	0,243	-0,071	0,224
13	721-840	-0,397	0,000	-0,273	0,007	-0,254	0,004	-0,220	0,027
14	781-900	-0,050	0,119	0,068	0,449	-0,105	0,136	-0,138	0,034
15	841-960	-0,010	0,903	-0,214	0,028	-0,007	0,935	-0,270	0,026
16	901-1020	0,025	0,213	-0,122	0,080	-0,005	0,955	-0,157	0,054
17	961-1080	0,095	0,379	0,070	0,142	-0,079	0,020	-0,071	0,132
18	1021-1140	-0,189	0,000	-0,192	0,000	-0,246	0,000	-0,170	0,000
19	1081-1200	-0,181	0,003	-0,013	0,897	-0,198	0,000	-0,181	0,003
20	1141-1260	-0,108	0,043	-0,128	0,041	0,043	0,759	-0,116	0,114
21	1201-1320	-0,316	0,034	-0,131	0,326	-0,279	0,119	-0,160	0,000
22	1261-1380	-0,348	0,000	-0,075	0,557	-0,144	0,077	-0,165	0,033
23	1321-1440	-0,165	0,000	-0,127	0,151	-0,165	0,028	-0,155	0,000
24	1381-1500	-0,101	0,244	-0,068	0,000	-0,211	0,016	0,030	0,847
25	1441-1560	-0,127	0,000	-0,285	0,000	-0,189	0,000	0,224	0,108
26	1501-1620	-0,123	0,028	-0,179	0,000	-0,123	0,007	-0,060	0,209
27	1561-1680	-0,675	0,001	0,054	0,000	-0,189	0,009	-0,210	0,025
28	1621-1740	-0,445	0,000	-0,208	0,002	-0,741	0,000	-0,348	0,000
29	1681-1800	-0,054	0,354	0,083	0,401	-0,064	0,339	-0,054	0,271
30	1741-1860	-0,187	0,202	-0,118	0,279	-0,049	0,651	-0,002	0,981
31	1801-1920	0,078	0,024	0,067	0,130	0,077	0,178	0,051	0,707
32	1861-1980	0,140	0,351	-0,192	0,253	-0,052	0,664	0,169	0,016

Zdroj: vlastní výpočty v EViews 6

Obrázek 2

Časový vývoj parametru β EGARCH modelu (4) pro C1

Zdroj: vlastní výpočty v EViews 6

Obrázek 3

Časový vývoj parametru γ EGARCH modelu (4) pro C1

Zdroj: vlastní výpočty v EViews 6

Obrázek 4

Časový vývoj podmíněné směrodatné odchylky σ_t EGARCH modelu (4) pro C1. Interval 421-540

Zdroj: vlastní výpočty v EViews 6

Obrázek 5

Časový vývoj podmíněné směrodatné odchylky σ_t EGARCH modelu (4) pro C1. Interval 1261-1380

Zdroj: vlastní výpočty v EViews 6

Výsledky provedené analýzy naznačují, že v obdobích předcházejících bodům zvratu ve sledované finanční řadě pravděpodobně nastávají výrazné změny modelových parametrů. Nejvýrazněji se to projevuje v chování denních relativních výnosů uzavíracích cen C1. V intervalu 1261-1380, kterému odpovídá časové období 05/02/2007-27/07/2007, byla odhadnutá hodnota parametru $\beta = -0,768$, což je absolutní minimum v celé časové řadě (obrázek 2). Jak již bylo zmíněno v úvodu, v tomto intervalu se rovněž nachází první maximum uzavírací ceny (hodnota $C = 1903,80$ v bodě 1366, tj. k datu 09/07/2007). Přitom byl parametr γ rovněž záporný a dosahoval hodnoty $\gamma = -0,348$ (obrázek 3).

Interpretace významu znaménka parametru β je zřejmá. Pro „obvyklé“ kladné hodnoty β lze očekávat „perzistentní“ chování podmíněného rozptylu, tj. relativně málo bodů zvratu v jeho grafickém vyjádření. Tato situace je znázorněna na obrázku 4 pro interval 421-540, tj. časové období 01/10/2003-26/03/2004, kdy docházelo k růstu uzavírací ceny a odhadnuté parametry byly $\beta = 0,950$ a $\gamma = 0,160$ (tabulka 1 a 2). Na obrázku 4 je pak znázorněn časový vývoj chování podmíněné směrodatné odchylky. Naproti tomu ve zmíněném intervalu 1261-1380 lze pozorovat „antiperzistentní“ chování podmíněné směrodatné odchylky s velkým počtem bodů zvratu (obrázek 5).

Dále byly odhadnuty parametry modelu EGARCH(1,1) pro celou časovou řadu denních výnosů uzavíracích cen indexu PX (celkem 1988 hodnot):

$$\hat{\omega} = -0,176773 \quad \hat{\beta} = 0,961888 \quad \hat{\alpha} = 0,260589 \quad \hat{\gamma} = -0,079953 \quad (5)$$

Dosadíme-li tyto odhady do pravé strany (4), dostaneme s přihlédnutím k (2)

$$\ln(\sigma_t^2) = \hat{\omega} + \hat{\beta} \ln(\sigma_{t-1}^2) + \hat{\alpha} |z_{t-1}| + \hat{\gamma} z_{t-1} \quad (6)$$

Odlogaritmováním tohoto vztahu získáme výraz pro podmíněný rozptyl

$$\sigma_t^2 = \exp(\hat{\omega}) \sigma_{t-1}^{2\hat{\beta}} \exp[\hat{\alpha}|z_{t-1}|] \exp[\hat{\gamma}z_{t-1}] \quad (7)$$

který udává závislost podmíněného rozptylu na velikosti a znaménku standardizovaného šoku v předcházejícím čase. Protože zpožděné hodnoty podmíněného rozptylu ovlivňují pouze násobnou konstantu v (7), lze je nahradit vhodně volenou střední hodnotou, nejčastěji mediánem M (tohoto způsobu používá i program Eviews 6). Asymetrickou reakci v závislosti na kladné a záporné šoky lze pak psát ve tvaru

$$\begin{aligned} \sigma_t^2 &= \exp(\hat{\omega}) M^{\hat{\beta}} \exp[(\gamma + \alpha)z_{t-1}] & \text{pro } z_{t-1} \geq 0 \\ \sigma_t^2 &= \exp(\hat{\omega}) M^{\hat{\beta}} \exp[(\gamma - \alpha)z_{t-1}] & \text{pro } z_{t-1} < 0 \end{aligned} \quad (8)$$

Dosažením odhadů (5) a hodnoty $M = 1,335899$ dostaneme konkrétní tvar závislosti

$$\begin{aligned} \sigma_t^2 &= 1,107155 \exp[0,180636 z_{t-1}] & \text{pro } z_{t-1} \geq 0 \\ \sigma_t^2 &= 1,107155 \exp[-0,34054 z_{t-1}] & \text{pro } z_{t-1} < 0 \end{aligned} \quad (9)$$

Na základě tohoto vztahu můžeme konstruovat graf NIC (News Impact Curve), znázorňující asymetrii reakce (Franses-vanDijk, 2006).

Obrázek 6

NIC křivka závislosti podmíněné směrodatné odchylky $\sigma_t(SIG2)$ na předchozím šoku $z_{t-1}(Z)$

3. Klouzavé charakteristiky

V důsledku heteroskedasticity a změn režimového chování je v případě finančních časových řad zřejmě nevhodné použití modelů s pevnými parametry. Naopak, použití klouzavých charakteristik může pomoci při identifikaci časových úseků, v nichž dochází ke strukturálním změnám (Dacorogna et al., 2001). Ve všech případech budeme uvažovat 21-denní délku klouzavé části, což odpovídá průměrnému měsíčnímu počtu obchodních dní.

Klouzavá směrodatná odchylka (Standard Deviation) vypočtená z minulých 21 hodnot denních relativních výnosů uzavíracích cen $C1$ je dána výrazem

$$SD_t(C1) = \sqrt{\frac{1}{n-1} \sum_{i=0}^{20} (C1_{t-i} - \overline{C1}_t)^2} \quad (10)$$

kde $\overline{C1}_t$ značí průměrný výnos v minulém 21-denním období.

Obrázek 7

Klouzavá směrodatná odchylka denních relativních výnosů uzavíracích cen indexu PX

Z obrázku 7 je patrné, že klouzavá směrodatná odchylka výnosů dosahuje minimální hodnoty v bodě 1448, tj. 3 dny po dosažení absolutního maxima (bod 1445).

Klouzavé cenové rozpětí zde definujeme v symbolickém tvaru

$$COHL_t = 100 \frac{MA_t(C) - MA_t(O)}{MA_t(H) - MA_t(L)} \quad (11)$$

kde operátor MA značí klouzavý průměr z 21 minulých hodnot příslušných cen. Tato charakteristika již využívá vnitrodenních změn kurzu.

Obrázek 8

Klouzavé cenové rozpětí indexu PX

Z obrázku 8 je patrný růst této charakteristiky až do bodu 1435, tj. 10 dnů před dosažením absolutního maxima v bodě 1445. Tento růst je pak vystřídán poklesovou fází až do záporných hodnot.

Klouzavé součty bodů zvratu se týkají časové řady denních výnosů uzavírací ceny a jsou počítány pro 21-denní období. Podmínka výskytu horního či dolního bodu zvratu je potom

$$(CI_{t-1} < CI_t > CI_{t+1}) \vee (CI_{t-1} > CI_t < CI_{t+1}) \quad (12)$$

Za předpokladu, že je časová řada realizací i.i.d. procesu, je počet bodů zvratu S v časovém intervalu délky n náhodnou veličinou se střední hodnotou a rozptylem (Cipra, 2008)

$$E(S) = \frac{2}{3}(n-2) \quad \text{var}(S) = \frac{16n-29}{90} \quad (13)$$

Pro dostatečně velká n se používá asymptotická verze testu náhodnosti s kritickým oborem

$$U = \frac{|S - E(S)|}{\sqrt{\text{var}(S)}} \geq u_{1-\alpha/2} \quad (14)$$

V našem případě je $n = 21$ a pro $\alpha = 0,05$ bude $u_{0,975} = 1,96$ a tedy

$$E(S) = 12,67 \quad \text{var}(S) = 3,41 \quad P(9,05 < S < 16,29) = 0,95 \quad (15)$$

Obrázek 9

Klouzavé součty bodů zvratu v časové řadě denních výnosů indexu PX

Z obrázku 9 je patrné, že před dosažením prvního i absolutního maxima je počet bodů zvratu výrazně vyšší než u ryze náhodného procesu (jedná se o tzv. antipersistenční chování).

4. Cenové indikátory technické analýzy

Technická analýza je metoda odhadu vývoje akciových cen, založená na studiu chování individuálních akcií a celkového trhu (Achelis, 2000; Víšková, 1997). Pravděpodobný budoucí vývoj cen je predikován na základě informací o minulých cenách a minulých objemech obchodů. Přitom se předpokládá, že

- Tržní ceny akcií obsahují všechny relevantní ekonomické, politické a psychologické informace. Aktuální ceny akcií jsou dány poměrem nabídky a poptávky
- Pohyby cen akcií vždy tvoří určitý trend, projevující se v určitém období
- Vzory tržního chování investorů se opakují; z toho důvodu lze usuzovat na budoucí cenový vývoj z minulého chování

Technický indikátor je funkce minulých akciových cen a minulých objemů obchodů, nabývající v každém obchodním dnu t určité hodnoty. Technické indikátory jsou součástí obchodních systémů a generují obchodní signály pro nákup (BUY) nebo prodej (SELL). Pro naše účely zde použijeme některé vybrané cenové indikátory a signály typu SELL, které budou odpovídat nástupu klesajícího trendu.

Jednoduchý klouzavý průměr – Simple Moving Average (SMA) je počítán jako jednoduchý klouzavý průměr z posledních n pozorování

$$SMA_t = \frac{1}{n} \sum_{i=0}^{n-1} X_{t-i} \quad (16)$$

Prodejní signál je v krátkodobém horizontu generován protnutím ceny a klouzavého průměru

$$(X_{t-1} \geq SMA_{t-1}) \wedge (X_t < SMA_t) \Rightarrow SELL \quad (17)$$

Zřejmou nevýhodou je však značný počet signálů. Proto je výhodnější použít průsečíky krátkodobého (délka n_1) a dlouhodobého (délka n_2) klouzavého průměru

$$\{SMA_{t-1}(n_1) \geq SMA_{t-1}(n_1)\} \wedge \{SMA_t(n_1) < SMA_t(n_2)\} \Rightarrow SELL \quad (18)$$

Z hlediska naší analýzy je přirozené volit $n_1 = 5$ (krátkodobé týdenní klouzavé průměry) a $n_2 = 21$ (dlouhodobé měsíční klouzavé průměry).

Obrázek 10

Indikátor SMA: průsečíky 5-denních a 21-denních klouzavých průměrů

Na obrázku 10 jsou vidět pozice prodejních signálů: první v bodě 1363, tj. 3 dny před dosažením prvního maxima, druhý v bodě 1445, tedy v koincidenci s absolutním maximem.

Bollingerova pásma – Bollinger Bands (BB) patří do skupiny pásmových (obálkových) indikátorů. V tomto případě je šířka pásma variabilní v závislosti na volatilitě akciových cen. Obecně lze horní pásmo (index U = Upper) a dolní pásmo (index L = Lower) konstruovat jako

$$BB_t^U = SMA_t(n) + c\sigma_t(n), \quad BB_t^L = SMA_t(n) - c\sigma_t(n) \quad (19)$$

s doporučenými hodnotami parametrů $n = 14$, $2 < c < 3$. Mírou volatility ceny je směrodatná odchylka $\sigma_t(n)$. Prodejní signál je pak dán protnutím horního pásma shora

$$(X_{t-1} > BB_{t-1}^U) \wedge (X_t > BB_t^U) \Rightarrow SELL \quad (20)$$

Prodejní signály jsou generovány v bodech 1369 a 1448, tedy vždy po 3 dnech od dosažení maxim (obrázek 11).

Obrázek 11

Indikátor Bollinger Bands s parametry $n = 20$, $c = 2,25$

Index relativní síly – Relative Strength Index (RSI) patří do skupiny oscilátorů. Je definován jako (Víšková, 1997)

$$RSI_t(n) = 100 - \frac{100}{1 + P(n)/N(n)} \quad 0 < RSI_t(n) < 100 \quad (21)$$

kde $P(n)$ a $N(n)$ jsou průměrné kladné a záporné změny cen za posledních n obchodních dnů

$$P(n) = \frac{1}{n} \sum_{i=0}^{n-1} (X_{t-i} - X_{t-i-1})^+ \quad N(n) = \frac{1}{n} \sum_{i=0}^{n-1} |(X_{t-i} - X_{t-i-1})|^- \quad (22)$$

Obrázek 12
Indikátor RSI s parametrem $n = 14$

Vysoké hodnoty RSI potom signalizují převládající kladné změny cen

$$P(n)/N(n) \rightarrow \infty \quad RSI_t(n) \rightarrow 100 \quad (23)$$

zatímco nízké hodnoty RSI signalizují převládající záporné změny cen

$$P(n)/N(n) \rightarrow 0 \quad RSI_t(n) \rightarrow 0 \quad (24)$$

Prodejní signály jsou generovány při překročení pevné horní hranice U s doporučením

$$RSI_t(n) > U \Rightarrow SELL \quad 70 \leq U \leq 80 \quad (25)$$

Z obrázku 12 je patrné, že horní hranice $U = 80$ je překročena v bodech 1429-1434, tedy 11-16 dnů před dosažením absolutního maxima.

Stochasik – Stochastic Oscillator (SO) využívá rovněž informace obsažené ve vnitrodenním cenovém pohybu. Jeho hodnoty se vypočtou ve dvou krocích (Víšková, 1997)

$$K_t(m, k) = 100 \frac{\sum_{j=0}^{k-1} (C_{t-j} - \min_i L_{t-i-j})}{\sum_{j=0}^{k-1} (\max_i H_{t-i-j} - \min_i L_{t-i-j})} \quad i = 0, 1, \dots, m-1 \quad (26)$$

$$SO_t(n) = SMA_t(n, K_t(m, k))$$

s doporučenými hodnotami $m = 21$, $k = n = 3$. Prodejní signál SELL bude generován, jestliže

$$SO_{t-1} \leq h \wedge SO_t > h \Rightarrow SELL \quad 70 < h < 90 \quad (27)$$

Při volbě $h = 90$ dostaneme předzvěsti zlomu trendu v bodě 1367 (1 den po prvním maximu) a v bodě 1427 (18 dnů před absolutním maximem).

Obrázek 13

Indikátor Stochastic Oscillator s parametry $m = 21$, $k = n = 3$

5. Umělé neuronové sítě

Umělé neuronové sítě (Artificial Neural Networks = ANN) jsou v současné době využívány v řadě aplikací, souvisejících s modelováním a predikcí (Bishop, 1995; Haykin, 2008; McNelis, 2005). Hlavní výhodou tohoto nástroje je možnost aproximace téměř jakékoli nelineární funkce s libovolnou přesností. Zejména ve finančních časových řadách s komplexními dynamickými nelineárními relacemi mohou být výhodnější než parametrické lineární modely. Na druhé straně je obtížné interpretovat význam parametrů a z toho důvodu jsou ANN často uvažovány jako „černá skříňka“, použitelná pro rozpoznávání obrazů a predikci.

ANN se obecně skládají z několika vrstev. *Vstupní vrstva* je tvořena individuálními vstupy, které hrají roli vysvětlujících proměnných. Tyto vstupy jsou násobeny intenzitami spojení – ve statistické terminologii *váhy*. Dále existuje jedna či více *skrytých vrstev*, z nichž každá obsahuje určitý počet *neuronů*. Ve skrytých vrstvách se tvoří lineární kombinace vstupů a je pak dále transformována *aktivační funkcí*. Nakonec se získá *výstup* jako vážený průměr transformovaných hodnot.

Obrázek 14

Schéma umělé neuronové sítě s jednou skrytou vrstvou obsahující dva neurony

Na obrázku 14 je schéma ANN se 3 vstupy a jednou skrytou vrstvou, obsahující 2 neurony. Matematický popis funkce ANN je výhodné provést ve třech krocích:

$$\begin{aligned}
 n_{k,t} &= w_{k,0} + \sum_{i=1}^I w_{k,i} x_{i,t} \\
 N_{k,t} &= 1 / [1 + \exp(-n_{k,t})] \\
 Y_t &= \gamma_0 + \sum_{k=1}^K \gamma_k N_{k,t} + \sum_{i=1}^I \beta_i x_{i,t}
 \end{aligned} \tag{28}$$

První rovnice popisuje vytvoření lineární kombinace vstupních proměnných, druhá rovnice pak jejich transformaci pomocí logsigmoidní aktivační funkce. Třetí rovnice vysvětluje, že výstupní hodnota může být získána buď jako lineární kombinace transformovaných hodnot nebo jako lineární kombinace vstupních proměnných. Neexistuje-li žádná vstupní vrstva, redukuje se model na vícenásobnou lineární regresi.

Jako příklad uvažujme systém popsaný následující soustavou rovnic

$$n_{1,t} = +40 + 10y_{t-1}, \quad n_{2,t} = y_{t-1}, \quad n_{3,t} = -20 + 20y_{t-1} \tag{29}$$

$$\begin{aligned}
 N_{1,t} &= 1 / [1 + \exp(-40 - 10y_{t-1})] \\
 N_{2,t} &= 1 / [1 + \exp(-y_{t-1})] \\
 N_{3,t} &= 1 / [1 + \exp(20 - 20y_{t-1})]
 \end{aligned} \tag{30}$$

$$Y_t = 2 + 8 N_{1,t} - 12 N_{2,t} + 6 N_{3,t} \tag{31}$$

Zřejmě se jedná o ANN s jednou vstupní proměnnou y_{t-1} a třemi neurony v jediné vstupní vrstvě. Graf výstupní proměnné Y_t v závislosti na vstupní proměnné y_{t-1} je následující

Obrázek 15

Výstupní signál Y_t tvořený umělou neuronovou sítí (21), (22), (23)

Z obrázku 15 je zřejmá silně zvýšená citlivost Y_t na y_{t-1} v okolí hodnot $y_{t-1} = -4$ a $y_{t-1} = 1$ a silně nelineární závislost, daná použitím sigmoidní aktivační funkce.

Pro vlastní analýzu byla vybrána časová řada týdenních hodnot uzavírací ceny indexu PX, celkem 421 hodnot. Velké množství denních hodnot totiž může vést k přetrénování sítě a snížení její predikční schopnosti. První maximum odpovídá bodu 274 (hodnota 1897,60 dosažená 04/07/2007), absolutní maximum pak bodu 290 (hodnota 1936,90 dosažená 29/10/2007). V prvním (resp. druhém) případě bylo použito 138 (resp. 148) hodnot pro trénink sítě, 68 (resp. 71) hodnot pro validaci a 68 (resp. 71) hodnot pro testování. V obou případech byly pomocí natrénované sítě provedeny predikce uzavírací ceny na 10 týdnů, tj. pro časy 275-284, resp. 291-300. Vysvětlující proměnnou tvořily zpožděné hodnoty uzavíracích cen se zpožděními $n = 5, 10$ a 20 týdnů, použité sítě pak obsahovaly buď jednu, nebo dvě skryté vrstvy s maximálně 8 neurony v každé vrstvě. Predikční schopnost jednotlivých sítí byla posuzována střední absolutní chybou předpovědí MAE (Mean Absolute Error). Výsledky jsou shrnuty v tabulce 3 a 4 a na obrázku 16 a 17. V tabulkách jsou pro každé zpoždění uvedeny 3 nejúspěšnější modely podle MAE. Na obrázcích je uváděn pouze nejúspěšnější model. Typ sítě je dán strukturou: např. 8-7 značí síť se dvěma skrytými vrstvami, 8 neuronů v první a 7 ve druhé vrstvě, zatímco 3 značí síť s jednou skrytou vrstvou obsahující 3 neurony.

Tabulka 3

Predikce uzavírací ceny indexu PX z prvního maxima (bod 274)

Bod	Close	n = 5			n = 10			n = 20		
Typ sítě		8-7	3	8-6	8-4	3	2	8-3	8-7	5
275	1863,40	1851,64	1821,57	1820,00	1833,41	1815,59	1820,18	1809,20	1854,57	1780,18
276	1854,90	1828,30	1781,79	1792,75	1792,63	1788,59	1786,00	1802,16	1851,11	1766,38
277	1771,40	1820,83	1759,61	1751,94	1795,91	1774,73	1763,47	1770,49	1831,19	1731,41
278	1772,30	1808,15	1747,07	1748,32	1772,80	1754,66	1745,89	1753,23	1802,24	1698,93
279	1750,60	1783,71	1726,45	1694,23	1750,88	1735,56	1726,65	1723,41	1792,38	1671,46
280	1688,10	1772,05	1713,46	1679,60	1749,20	1716,02	1693,93	1715,44	1756,91	1648,77
281	1767,10	1762,54	1704,65	1656,14	1753,21	1711,08	1674,76	1685,62	1741,66	1619,09
282	1789,30	1750,53	1697,68	1652,69	1726,64	1699,21	1664,75	1683,10	1688,31	1593,43
283	1762,40	1740,39	1691,06	1630,34	1731,33	1696,11	1635,16	1670,14	1677,56	1579,53
284	1761,90	1733,77	1686,16	1629,69	1742,64	1694,16	1621,15	1666,38	1624,25	1552,54
MAE		33,42	50,26	72,57	30,55	45,82	66,11	55,69	56,18	113,97

Tabulka 4

Predikce uzavírací ceny indexu PX z absolutního maxima (bod 290)

Bod	Close	n = 5			n = 10			n = 20		
Typ sítě		6	5	8-6	5-3	4	8-8	5	8-6	3-3
291	1863,40	1865,33	1828,00	1901,76	1858,97	1845,24	1879,34	1851,08	1860,54	1867,27
292	1854,90	1851,81	1781,40	1909,31	1840,50	1835,68	1859,51	1814,81	1830,34	1855,27
293	1771,40	1836,21	1750,85	1914,26	1816,38	1796,48	1829,69	1795,21	1830,46	1842,89
294	1772,30	1823,86	1726,09	1909,67	1778,49	1767,09	1835,06	1780,31	1808,61	1836,57
295	1750,60	1809,17	1698,62	1912,11	1782,19	1748,74	1819,22	1779,91	1783,59	1838,26
296	1688,10	1803,23	1685,10	1911,81	1756,86	1720,57	1819,13	1776,38	1834,44	1834,04
297	1767,10	1796,79	1675,04	1911,65	1761,68	1711,12	1791,53	1769,19	1829,88	1825,40
298	1789,30	1792,35	1667,37	1912,03	1757,68	1699,99	1797,37	1768,99	1812,05	1824,80
299	1762,40	1789,16	1662,17	1911,97	1749,36	1692,11	1795,58	1763,83	1839,29	1817,08
300	1761,90	1787,05	1659,30	1912,06	1763,68	1695,98	1788,32	1770,58	1835,32	1811,45
MAE		51,80	72,20	126,63	42,39	49,84	50,80	37,89	53,54	59,89

Obrázek 16

Predikce uzavírací ceny indexu PX z prvního maxima (bod 274): model 8-4, n=10

Plná čára se čtverci = skutečné hodnoty

Přerušovaná čára s trojúhelníky = predikované hodnoty

Obrázek 17

Predikce uzavírací ceny indexu PX z absolutního maxima (bod 290): model 5, n=20

Plná čára se čtverci = skutečné hodnoty

Přerušovaná čára s trojúhelníky = predikované hodnoty

Z dosažených výsledků plyne, že nelze jednoznačně rozhodnout, zda je lepší použití ANN s jednou nebo dvěma skrytými vrstvami. Ani závislost výsledků na počtu zpoždění vysvětlující proměnné není zřejmá. Např. predikce z prvního maxima byla vždy nejúspěšnější při použití modelů se dvěma skrytými vrstvami a lepší výsledky byly dosaženy při zpožděních $n = 10$ a $n = 5$, kdy byly hodnoty MAE zhruba poloviční než pro $n = 20$.

Naopak, při predikci z absolutního maxima byly ve dvou případech úspěšnější modely s jednou skrytou vrstvou a s delším zpožděním. Je třeba dodat, že lineární

modely bez skrytých vrstev z predikčního hlediska naprosto selhaly a jejich použití by vedlo dokonce k předpovědím rostoucího trendu namísto klesajícího; proto zde tyto výsledky neuvádíme.

6. Závěr

Z provedených analýz vyplývá, že predikce změny směru trendu v časových řadách vyžaduje multidisciplinární přístup a kombinaci různých postupů. Studium chování různých klouzavých charakteristik může pomoci při identifikaci úseků časové řady s různým režimovým chováním. Cennou informací zřejmě přinášejí také vnitrodenní variace cen (otevírací, uzavírací, minimální, maximální), které zřejmě odrážejí převládající tendence v počtu prodejních a nákupních příkazů.

Výsledky získané aplikací klouzavých modelů EGARCH(1,1) naznačují, že v obdobích předcházejících bodům zvratu ve sledované finanční řadě pravděpodobně nastávají výrazné změny modelových parametrů. Nejvýrazněji se to projevuje v chování denních relativních výnosů uzavíracích cen C1. Konkrétně se jedná o fakt, že záporné hodnoty parametru β mohou signalizovat změnu rostoucího trendu na klesající, zejména pak v konjunkci se zápornými hodnotami parametru γ . Interpretace významu znaménka parametru β je zřejmá. Pro „obvyklé“ kladné hodnoty β lze očekávat „perzistentní“ chování podmíněného rozptylu, tj. relativně málo bodů zvratu v jeho grafickém vyjádření. Naproti tomu v období před dosažením maximální hodnoty uzavírací ceny lze pozorovat „antiperzistentní“ chování podmíněné směrodatné odchylky s velkým počtem bodů zvratu.

Metody technické analýzy se již staly standardní součástí procedur ke stanovení obchodních signálů. Jak však správně upozornili Franes a vanDijk (2006), tyto metody se zpočátku setkávaly s jistým skepticismem a teprve v posledních letech došlo k jejich uznání na základě řady provedených studií (Franes–vanGriensven, 1998). V naší analýze byla prokázána dobrá predikční schopnost několika vybraných indikátorů; přesvědčivé závěry však bude možno provést až po studiu více časových řad pomocí indikátorů s optimalizovanými hodnotami parametrů.

Použití umělých neuronových sítí (ANN) vedlo k nadějným výsledkům z hlediska predikce v 10-týdenním horizontu, i zde je však problémem volba struktury sítě a počet zpožděných hodnot. Je třeba si uvědomit, že jak technická analýza, tak ANN patří do oblasti umělé inteligence a jedná se o procesy učení z minulého chování a rozpoznávání obrazů.

Závěrem je třeba poznamenat, že dosažené výsledky bude nutno dále ověřovat na dalších časových řadách kapitálového trhu z různých světových oblastí a z různých časových období, což bude předmětem dalšího zkoumání.

Literatura

- ACHELIS, S. 2000. *Technical Analysis from A to Z*. New York: McGraw-Hill, 2000. ISBN 978-0071363488.
- BISHOP, C. 1995. *Neural Networks for Pattern Recognition*. Oxford: Oxford University Press, 1995. ISBN 0-19-853864-2.
- BOLLERSLEV, T. 1986. Generalized Autoregressive Conditional Heteroskedasticity. *Journal of Econometrics*. 1986, Vol. 31, pp. 307–327.
- CIPRA, T. 2008. *Finanční ekonometrie*. Praha: Ekopress, 2008. ISBN 978-80-86929-43-9.
- ENGLE, R. 1982. Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of United Kingdom Inflation. *Econometrica*. 1982, Vol. 50, pp. 987–1007.
- FRANSES, P.; vanDIJK, D. 2006. *Nonlinear Time Series Models in Empirical Finance*. Cambridge: Cambridge University Press, 2006. ISBN 0-521-779650.
- FRANSES, P.; vanGRIENSVEN, K. 1998. Forecasting exchange rates using neural networks for technical trading rules. *Studies in Nonlinear Dynamics and Econometrics*. 1998, Vol. 2, pp. 109–116.
- GENCAY, R.; STENGOS, T. 1998. Moving average rules, volume and the predictability of security returns with feedforward networks. *Journal of Forecasting*. 1998, Vol. 17, No. 5-6, pp. 401–414.
- HAYKIN, S. 1999. *Neural Networks: A Comprehensive Foundation*. Upper Saddle River: Prentice Hall, 1999. ISBN 0-13-273350-1.
- KEDEM, B.; FOKIANOS, K. 2002. *Regression Models for Time Series Analysis*. New Jersey: Wiley, 2002. ISBN 0-471-36355-3.
- McNELIS, P. 2005. *Neural Networks in Finance*. Boston: Elsevier Academic Press, 2005. ISBN 0-12-485967-4.
- NELSON, D. B. 1991. Conditional Heteroskedasticity in Asset Returns: A New Approach. *Econometrica*. 1991, Vol. 59, pp. 347–370.
- VÍŠKOVÁ, H. 1997. *Technická analýza akcií*. Praha: HZ, 1997. ISBN 80-86009-13-0.

SELECTED METHODS OF THE PREDICTION OF PX INDEX TREND REVERSAL

Jiří Trešl, University of Economics, Prague, W. Churchill Sq. 4, CZ – 130 67 Prague 3
(tresl@vse.cz)

Abstract

The paper is concerned with the use of several methods that can be useful from the point of view of trend reversal in financial time series. These methods are demonstrated on PX index time series during 2002-2009. The research itself is subdivided into four parts corresponding to individual analytical methods used. The first group contains the use of moving EGARCH(1,1) model to daily relative returns of PX index. The results obtained are indicative of the importance of negative parameter values, which can be considered as precursors of the trend reversal. The second group contains different moving characteristics that are able to signalize regime changes in certain time intervals. Particularly, the information related to intraday price variations proved to be useful. Third, selected price indicators from technical analysis were employed. Among them, Simple Moving Averages, Bollinger Bands, Relative Strength Index and Stochastic led to acceptable predictions. Last, the predictive ability of Artificial Neural Networks was tested with respect to different network structure and number of delayed values of explanatory variable. The results obtained here are promising, but further research in this direction is necessary.

Keywords

financial time series, trend reversal, technical analysis, artificial neural networks

JEL Classification

C22, C45, C53