

UNIVERZITA KOMENSKÉHO V BRATISLAVE
FAKULTA MATEMATIKY, FYZIKY A INFORMATIKY



**Predpovedanie finančných časových radov pomocou
neurónových sietí**

Diplomová práca

2009

Jana Orsagová

Predpovedanie finančných časových radov pomocou neurónových sietí

DIPLOMOVÁ PRÁCA

Diplomant: Jana Orsaghová

UNIVERZITA KOMENSKÉHO V BRATISLAVE
FAKULTA MATEMATIKY, FYZIKY A INFORMATIKY
KATEDRA APLIKOVANEJ MATEMATIKY A ŠTATISTIKY
EKONOMICKÁ A FINANČNÁ MATEMATIKA

Vedúci práce: Doc. Mgr. Marian Grendár, PhD.

BRATISLAVA 2009

Čestné prehlásenie

Čestne prehlasujem, že som túto diplomovú prácu vypracovala samostatne, len s pomocou nadobudnutých teoretických vedomostí, konzultácií a uvedenej literatúry.

V Bratislave, apríl 2009

.....

Jana Orsaghová

Pod'akovanie

Touto cestou chcem pod'akovať všetkým, ktorí mi akýmkoľvek spôsobom pomohli pri písaní diplomovej práce. Veľmi pekne ďakujem svojmu vedúcemu Marianovi Grendárovi za odborné vedenie, pripomienky, návrhy a čas, ktorý mi venoval pri vypracovávaní diplomovej práce.

Abstrakt

Jana Orsaghová: *Predpovedanie finančných časových radov pomocou neurónových sietí*. [diplomová práca]. Univerzita Komenského v Bratislave. Fakulta matematiky, fyziky a informatiky. Katedra aplikovanej matematiky a štatistiky. Vedúci diplomovej práce: Marian Grendár. Bratislava. 2009. 55 strán

Jedným zo základných cieľov analýzy finančných časových radov je predpovedanie ich budúceho vývoja. Keďže ide o dáta vyznačujúce sa stochastickým správaním, je ich analýza pomerne zložitá. Na ich predpovedanie sa používa viacero metód, koncept neurónových sietí je jednou z nich. Doterajšie výskumy ukázali, že neurónové siete môžu byť pomerne úspešné. Dokážu zachytiť znaky a nepravidelnosti v dátach, pričom popisujú nielen lineárne ale aj nelineárne vzťahy medzi nimi. Táto ich schopnosť je nesmierne dôležitá pri modelovaní dynamického systému, akými finančné časové rady určite sú. V práci sa budeme zameriavať na praktickú stránku predpovedania, teoretické podchytenie neurónových sietí spomenieme iba stručne. Naším cieľom bude zanalyzovať a zhodnotiť možnosti využitia neurónových sietí na predpovedanie časových radov, konkrétne ich schopnosť predpovedať výmenné kurzy. Bude nás zaujímať nielen presnosť predikcie ale hlavne ziskovosť, ktorú môžeme získať trhovou simuláciou obchodnej stratégie určenej na základe predpovedaných hodnôt.

Kľúčové slová: analýza časových radov, dopredné neurónové siete, metóda spätného šírenia chyby

Obsah

Úvod	3
1. Klasické metódy predpovedania finančných časových radov	6
1.1 Predpovedateľnosť	6
1.2. Box-Jenkinsova metodológia	7
1.3. Modely stochastickej volatility ARCH a GARCH	8
2. Neurónové siete	10
2.1. Neurón	10
2.2. Neurónové siete	12
2.3. Neurónové siete vyššieho rádu	13
2.4. Učenie	14
3. Učenie na základe korekcie chyby	15
3.1. Metódy riešenia úloh na voľný extrém	15
3.1.1. Gradientová metóda najväčšieho spádu	16
3.1.2. Newtonova metóda	16
3.1.3. Levenberg-Marquardtova metóda	17
3.2. Delta pravidlo	18
3.3. Metóda spätného šírenia chyby	18
4. Meranie výkonnosti	21
5. Dáta	23
5.1. Štatistické vlastnosti dát	24
6. Štúdia predikčnej schopnosti neurónových sietí	27
6.1. Návrh siete	27
6.1.1. Príprava dát	27
6.1.2. Výber architektúry	29
6.1.3. Trénovací algoritmus	30
6.1.4. Zastavenie tréningu	30
6.2. Analýza predikčnej schopnosti neurónových sietí	32

6.2.1. Predpovedanie výmenných kurzov	32
6.2.2. Predpovedanie výnosov výmenných kurzov	34
6.3. Porovnanie klasických neurónových sietí a neurónových sietí druhého rádu	35
6.3.1. Porovnanie pre výmenné kurzy	35
6.3.2. Porovnanie pre výnosy výmenných kurzov	36
6.4. Porovnanie predikčnej schopnosti neurónových sietí s inými metódami	38
6.4.1. Porovnanie predpovede s ARMA - GARCH modelom	38
6.4.2. Porovnanie kvality predpovede s alternatívnymi metódami	41
Záver	43
Literatúra	46
Príloha A : Hurstov koeficient a R/S analýza	49
Príloha B : Odvodenie metódy spätného šírenia chyby	51
Príloha C : Predpovedanie volatility	55

Úvod

Výmenné kurzy patria ku významným aktívam obchodovaným na medzinárodných trhoch. Už nejaký čas sa dostávajú do centra pozornosti vedcov, ktorí sa zaujímajú o vysvetlenie ich pohybov. Tento záujem bol podnietený pádom Breton-Woodskeho systému v roku 1973, kedy sa mnohé industrializované ekonomiky zriekli fixných výmenných kurzov a prešli k plávajúcim. Existuje množstvo motivácií na predpovedanie kurzu. Tou najzákladnejšou je finančný zisk. Akýkoľvek systém, ktorý by dokázal predpovedať čo i len nárast resp. pokles by priniesol svojmu majiteľovi vysoký zisk. Predpovedanie výmenného kurzu je však veľmi zložitý. Je to podmienené zložitým volatilným správaním a vysokým šumom, čo je dôsledkom zložitého trhového mechanizmu generujúceho dáta. Preto sa viacerí domnievajú, že finančné časové rady sú dokonca nepredvídateľné.

Hypotéza efektívneho trhu (EMH, z angl. *Efficient Market Hypothesis*) hovorí, že súčasná cena aktív odzrkadľuje všetky dostupné informácie v danom čase. Informácie ovplyvňujúce budúce ceny aktív nemáme v súčasnosti k dispozícii, pričom EMH ich považuje za náhodné. Preto nie je možné predpovedať budúci vývoj cien aktíva na základe jeho historických hodnôt alebo historických hodnôt iných finančných ukazovateľov. V skutočnosti však trh nereaguje na nové informácie okamžite ale s oneskorením. Významnú úlohu pri tom zohráva psychologický faktor. Toto podnietilo úsilie mnohých výskumníkov o vyvrátenie tejto hypotézy – snažili sa ukázať, že trhy nie sú efektívne. V súčasnosti neexistuje konsenzus medzi prívržencami a odporcami EMH. Snaha predpovedať ceny aktív však kvôli ekonomickému zisku neustáva.

Istým variantom ku EMH je teória náhodnej prechádzky (RWT z angl. *Random Walk Hypothesis*). Podľa nej sledujú ceny aktív náhodný pohyb a preto sú navzájom nezávislé. Keďže cena aktív s rovnakou pravdepodobnosťou vzrastie i klesne, akákoľvek snaha predpovedať trh je márna. Podľa niektorých jej zástancov je najlepšou investičnou možnosťou Buy & Hold stratégia. Výskumy však poukázali na existenciu viacerých radov, ktoré nemajú náhodné správanie. Trhy nie sú úplne efektívne a preto v dátach môžeme nájsť trendy a korelácie.

Existujú dva základné prístupy ku modelovaniu finančných časových radov a to technická a fundamentálna analýza. Fundamentálna analýza používa na modelovanie konkrétneho typu dát iné ekonomické dáta, ktoré môžu nejakým spôsobom súvisieť s našim radom dát. Tým do istej miery objasňuje ekonomický kontext jednotlivých predpovedí. V tejto práci sa však pustíme opačnou cestou, teda technickou analýzou. Tá kladie dôraz čisto na využitie informácií z historických dát a neberie do úvahy žiadne ekonomické faktory.

Finančné časové rady sa vyznačujú vlastnosťami, ktoré komplikujú ich predpovedanie. Sú to hlavne nestacionarita, nelinearita a zašumenosť. Klasické štatistické metódy sa bežne

používajú pri technickej analýze. Napríklad ARMA modely (z angl. *autoregressive moving average model*) sa ukázali ako efektívny nástroj na modelovanie niektorých finančných časových radov. Ide však iba o lineárne modely, ktoré nie sú schopné vyčítať z dát nelineárne vzťahy. Preto rastie záujem o nelineárne modely, ktoré by boli schopné zachytiť nelineárne správanie finančných časových radov. Medzi najznámejšie patria ARV modely (z angl. *autoregressive random variance model*) a ARCH modely (z angl. *autoregressive conditional heteroskedasticity model*). K nelineárnym modelom patria i neurónové siete. Vďaka svojmu neparametrickému adaptívnemu správaniu a tolerancii voči šumu dokážu popísať dynamiku nestacionárnych časových radov. Sú veľmi flexibilné, preto sa dajú použiť na technickú i fundamentálnu analýzu.

Neurónové siete majú viaceré vlastnosti, ktoré ich robia atraktívnymi na predpovedanie finančných časových radov. Jedna z ich hlavných výhod v porovnaní so štatistickými metódami spočíva v tom, že nepotrebujú žiadne predpoklady o modeloch, stačia im iba vstupné dáta. Preto sú vhodné aj pre úlohy, ktorých riešenie by si pri iných metódach vyžadovalo dodatočné predpoklady resp. špecifikácie modelu. Navyše si dokážu poradiť s nízkou kvalitou dát. Ďalšou ich výhodou je schopnosť zovšeobecňovať. Sú univerzálnym aproximátorom, keďže sú schopné aproximovať ľubovoľnú funkciu do ľubovoľného stupňa presnosti.

Napriek dobrým výsledkom, ktoré dosahujú pri predpovedaní finančných časových radov, čelia kritike kvôli viacerým nedostatkom. Hlavný problém neurónových sietí spočíva vo voľbe správnej architektúry a výbere efektívneho tréningového algoritmu. Rôzne algoritmy neurónových sietí môžu na tých istých dátach produkovať rozdielne výsledky. Navyše i váhy a prahy neurónov sa môžu trénovať viacerými metódami.

Doterajšie výskumy ukázali, že neurónové siete môžu byť silným nástrojom používaným vo financiách. Používajú sa v oblastiach, kde sa využíva semiparametrická alebo neparametrická regresia prípadne klasifikácia parametrov, ako napríklad predpovedanie finančných časových radov, odhad rizika a skóringové modely. Zo štatistického pohľadu sú neurónové siete neparametrickými nelineárnymi regresnými modelmi. Matematický koncept neurónových sietí je však pomerne zložitý a nie všetky teoretické aspekty sú dotiahnuté do finálnej podoby. Veľa poznatkov o neurónových sieťach bolo získaných heuristickým prístupom ich aplikovaním na reálne problémy.

Cieľom tejto práce je zhodnotiť, či sa dajú predpovede získané neurónovými sieťami úspešne použiť pri obchodovaní na trhu. Pojednáme o teoretickom podchytení neurónových sietí a zodpovedajúcom matematickom popise. Popíšeme základný koncept dopredných neurónových sietí tréňovaných metódou spätného šírenia chyby a poznatky potrebné na ich úspešnú implementáciu. Povieme si o neurónových sieťach vyššieho rádu, ktoré sú schopné simulovať vyšší stupeň nelinearity dát. Budeme predpovedať nielen samotné výmenné kurzy, ale aj ich výnosy a volatilitu. Budeme experimentovať s dátami a nastavením siete s cieľom zistiť, akú predikčnú schopnosť dokážeme vytvoriť. Zaujímať nás bude nielen jej matematické vyjadrenie v absolútnych číslach ale ziskovosť na finančných trhoch, ktorú dosiahneme reálnymi simuláciami predpovedí. Na zhodnotenie ich predikčnej schopnosti

použijeme klasické štandardy. Budeme sa snažiť dopátrať k odpovedi, či sa neurónové siete dajú trénovať tak, aby dospeli ku spoľahlivým predpovediam budúceho vývoja.

Práca je rozčlenená do šiestich častí. Prvá kapitola obsahuje prehľad klasických štatistických metód používaných na analýzu finančných časových radov ako Box-Jenkinsova metodológia a modely stochastickej volatility. Druhá kapitola popisuje základné elementy a klasifikáciu neurónových sietí. Tretia kapitola vysvetľuje princíp učenia na základe korekcie chyby, pričom sa zameriava na metódu spätného šírenia chyby. Vo štvrtej kapitole zadefinujeme miery, ktorými budeme merať výkonnosť neurónovej siete. Piata kapitola vyhodnocuje štatistické vlastnosti vybraných finančných časových radov. Šiesta kapitola popisuje nastavenie neurónovej siete, dizajn architektúry, výber vstupno-výstupných párov, špecifikáciu parametrov siete. Prezентuje výsledky predpovedí a ich analýzu. Zhodnocuje schopnosti získaných predpovedí pri reálnej trhovej simulácii. Výsledky sú porovnané s klasickou Buy & Hold stratégiou, naivnou predikciou a v neposlednom rade s výsledkami dosiahnutými klasickými štatistickými metódami. Záver obsahuje zhrnutie a interpretáciu výsledkov, návrh možných zlepšení. V prílohe je bližšie teoretické pojednanie o témach spomenutých v práci.

Kapitola 1

Klasické metódy predpovedania finančných časových radov

Ako sme už v úvode spomenuli, v tejto práci sa budeme zaoberať takými modelmi, ktoré modelujú časové rady iba pomocou informácií obsiahnutých v ich vlastných minulých hodnotách. Vo všeobecnosti môžeme uvažované modely rozdeliť do dvoch kategórií a to na lineárne a nelineárne. Keďže my sa v tejto práci budeme zaoberať analýzou výmenných kurzov, budeme potrebovať modely schopné vysvetliť ich stochastické správanie a vysokú volatilitu. V nasledujúcich častiach sa preto bližšie pozrieme na Box-Jenkinsovu metodológiu a modely stochastickej volatility (Engle a Bollerslev).

1.1 Predpovedateľnosť

Prvá otázka, ktorú si treba položiť pri predpovedaní časových radov, je, či je uvažovaný časový rad vôbec predpovedateľný. Ak je časový rad náhodný, akýkoľvek pokus o predpoveď pravdepodobne zlyhá. Preto budeme chcieť identifikovať, či sú nami uvažované časové rady aspoň do istej miery predpovedateľné.

Hurstov koeficient predstavuje akýsi numerický odhad predpovedateľnosti časového radu. Vyhodnocuje prítomnosť pamäťových znakov v časovom rade dát. Časové rady s veľkou hodnotou Hurstovho koeficientu majú silný trend, takže je prirodzené predpokladať, že takéto časové rady sú skôr predpovedateľné ako časové rady s malým Hurstovým koeficientom [23]. Preto nás bude zaujímať výška Hurstovho koeficientu pre nami uvažovaný časový rad. Vypočítame ho pomocou neparametrickej štatistickej metódy nazývanej škálovanie rozsahu alebo R/S analýza (z angl. *rescaled range analysis*). Tá poskytuje nasledovnú aproximáciu Hurstovho koeficientu

$$(R/S)_n \approx c n^H,$$

kde n je dĺžka časového okna, H je Hurstov koeficient, $(R/S)_n$ je bezrozmerný pomer priemerného rozsahu a štandardnej odchýlky denných fluktuácií. Postup výpočtu Hurstovho koeficientu spolu s interpretáciou dosiahnutých hodnôt je popísaný v prílohe A.

Ďalšia metóda, ktorou by sa dala preskúmať povaha uvažovaného časového radu, vychádza z hodnôt autokorelačnej funkcie. Ak autokorelácia klesá veľmi pomaly, ide o proces s dlhodobou pamäťou (z angl. *long memory process*).

1.2. Box-Jenkinsova metodológia

V praxi sa často stretávame s lineárnymi stochastickými modelmi časových radov, ktoré sú známe pod názvom Box-Jenkinsova metodológia. Existujú tri základné typy lineárnych modelov: AR modely (z angl. *Autoregressive model*), MA modely (z angl. *Moving average model*) a ARMA modely (z angl. *Autoregressive moving average model*).

Stacionarita radov

Kľúčovým pojmom Box-Jenkinsovej metodológie je stacionarita stochastického procesu. Hovoríme, že stochastický proces je striktnie stacionárny, ak je jeho pravdepodobnostné rozdelenie invariantné voči posunom v čase. Pretože sa táto vlastnosť iba ťažko overuje, zaviedol sa pojem slabej stacionarity stochastických procesov. Proces je slabo stacionárny, ak sú jeho stredná hodnota a disperzia invariantné voči posunom v čase. Akýkoľvek silno stacionárny proces so strednou hodnotou μ a disperziou σ^2 je tiež slabo stacionárny.

Väčšina finančných časových radov je nestacionárna, hoci modely prezentované v tejto časti sú stacionárne. Nestacionárne časové rady však vieme transformovať na stacionárne, preto môžu byť i tieto modely použité na analýzu nestacionárnych dát. Existuje viacero metód, ako to dosiahnuť. Pri finančných časových radoch sa často stretávame s transformáciou na výnosy, ktoré umožňujú dobrú ekonomickú interpretáciu získaných výsledkov. Na posúdenie, či je proces stacionárny, sa používajú autokorelačné funkcie a štatistické testy stacionarity, napr. rozšírený Dickey-Fullerov test. Autokorelačná funkcia je časový rad autokorelačných koeficientov, ktoré merajú stupeň korelácie medzi hodnotami x_t a x_{t-k} daného časového radu a tým udávajú informácie o sile lineárnej závislosti medzi nimi. Autokorelačný koeficient je definovaný vzťahom

$$\rho_t = \frac{\text{Cov}(x_{t-k}, x_t)}{\sqrt{D(x_{t-k})D(x_t)}}.$$

Keďže pravdepodobnostné rozdelenie dát nepoznáme, teoretický autokorelačný koeficient nevieme určiť. Môžeme však odhadnúť výberový autokorelačný koeficient r_t ,

$$r_t = \frac{\sum (x_t - \bar{x})(x_{t+k} - \bar{x})}{\sum (x_t - \bar{x})^2}.$$

Stacionaritu môžeme určiť posúdením grafov autokorelačnej funkcie. V prípade, že hodnoty korelačného koeficientu presahujú rozmedzie určené intervalmi spoľahlivosti iba pri malom počte omeškaní, rad je stacionárny. Šírka 95% intervalov spoľahlivosti je približne $\pm 2/\sqrt{N}$, kde N je rozmer dát.

Zo štatistických testov spomenieme rozšírený Dickey-Fuller test (skrátene ADF z angl. *Augmented Dickey-Fuller test*), ktorý testuje prítomnosť jednotkového koreňa (z angl. *unit root*) v autoregresnom modeli. Nech q je riešením charakteristickej rovnice príslušného modelu. Ak $|q| \geq 1$, rad je nestacionárny, navyše jeho variancia rastie s časom. V prípade, že $|q| < 1$, rad je stacionárny. Za predpokladu nulovej hypotézy jednotkového koreňa ADF štatistika nemá Studentovo t- rozdelenie. Pri testovaní nulovej hypotézy treba zvážiť voľbu

deterministických členov (trend a priesečník) a počtu omeškaní. To môžeme určiť pomocou informačných kritérií (Akaike, Schwarz, Hannan-Quinn).

ARMA modely

Všeobecnou triedou lineárnych modelov používaných na predpovedanie časových radov je trieda ARMA modelov. ARMA(p, q) model má tvar

$$y_t = \gamma + \sum_{i=1}^p \varphi_i y_{t-i} + \sum_{j=1}^q \theta_j e_{t-j} + e_t,$$

kde φ_i sú autoregresné parametre a θ_j sú parametre kľazových priemerov. V prípade, že proces y_t má nulovú strednú hodnotu, je γ rovné nule. Rezídua e_t sú nezávislé a rovnomerne rozdelené s vlastnosťami

$$E(e_t) = 0 \quad E(e_t^2) = \rho_e^2 \quad \text{Cov}(e_t, e_s) = 0 \quad s \neq t.$$

V prípade AR a MA procesov sa veľkosť p respektíve q určuje vizuálnym posúdením korelogramov autokorelačnej a parciálnej autokorelačnej funkcie. Pri ARMA procese je to však pomerne zložitá. Preto sa výška týchto koeficientov určuje porovnaním informačných kritérií pre viaceré hodnoty p a q .

Na odhad parametrov ARMA modelu je možné použiť viacero metód, napr. Yule-Walkerove vzorce alebo Burgov algoritmus. Podľa teórie predpovedania je optimálnou hodnotou predikcie tá, ktorá minimalizuje strednú kvadratickú chybu. Najlepším odhadom y_t je podmienená stredná hodnota v predchádzajúcich časových okamihoch, teda

$$\hat{y}_t = E(y_t | y_{t-1}, y_{t-2}, \dots).$$

Bolo dokázané [2], že aproximáciou optimálnej predpovede v prípade ARMA(p, q) modelov je rekurzívny vzťah

$$\hat{y}_t = \sum_{i=1}^p \varphi_i y_{t-i} + \sum_{j=1}^q \theta_j \hat{e}_{t-j}.$$

1.3. Modely stochastickej volatility ARCH a GARCH

Analýza finančných časových radov, ako sú ceny akcií, úrokové miery či výmenné kurzy, je pomerne zložitá, pretože tieto veličiny vykazujú vysokú volatilitu. V ekonometrii a financiách sú na vzostupe práve také modely, ktoré modelujú výnosy pomocou volatility. Medzi tieto modely patria aj ARCH (z angl. *autoregressive conditional heteroscedasticity*) a GARCH (z angl. *generalized autoregressive conditional heteroscedasticity*). Kľúčový prínos týchto modelov spočíva v rozlíšení podmienenej a nepodmienenej volatility procesu rezíduí $\{\varepsilon_t\}$. Pojem podmienenej volatility vyjadruje explicitnú závislosť na minulých pozorovaniach, zatiaľ čo nepodmienená volatilita súvisí s dlhodobým správaním sa časového radu.

ARCH predstavuje akronymom pre autoregresne podmienenú heteroskedasticitu. Ak je podmienená volatilita stochastického procesu heteroskedastická, ARCH model rádu q môžeme zapísať v tvare

$$\begin{aligned}y_t &= \mu + \varepsilon_t \\ \varepsilon_t &= z_t \sigma_t \\ \sigma_t^2 &= \omega + \alpha_1 \varepsilon_{t-1}^2 + \alpha_2 \varepsilon_{t-2}^2 + \dots + \alpha_q \varepsilon_{t-q}^2,\end{aligned}$$

kde $\omega > 0$, $\alpha_i \geq 0 \quad \forall i$, z_t je nezávisle a rovnomerne rozdelená náhodná premenná so špecifikovaným rozdelením. V štandardnom ARCH modeli je z_t zo štandardizovaného normálneho rozdelenia. Pri ARCH modeloch sa stretávame s veľkými hodnotami q a teda s veľkým počtom parametrov, ktoré je potrebné odhadovať. Preto bol vytvorený úspornejší model

$$\begin{aligned}y_t &= \mu + \varepsilon_t \\ \varepsilon_t &= z_t \sigma_t \\ \sigma_t^2 &= \omega + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^p \beta_j \sigma_{t-j}^2,\end{aligned}$$

pričom musí platiť

$$\sum_{i=1}^q \alpha_i + \sum_{j=1}^p \beta_j < 1 \quad \omega > 0 \quad \alpha_i \geq 0 \quad \forall i \quad \beta_j \geq 0 \quad \forall j.$$

Prvá podmienka musí byť splnená z dôvodu zabezpečenia jednoznačnosti a stacionarity, zvyšné tri kvôli kladným hodnotám volatility. Týmito rovnicami je definovaný zovšeobecnený ARCH(p,q) model, tiež známy pod názvom GARCH(p,q). V praxi sa stretávame iba s malými hodnotami parametrov p a q . Najčastejšie sa vyskytujú modely GARCH(1,1), GARCH(2,1) alebo GARCH(1,2). Už jednoduchý GARCH(1,1) s tromi parametrami v rovnici podmienenej volatility dokáže zachytiť väčšinu variability výnosov. Má tvar

$$\begin{aligned}y_t &= \mu + \varepsilon_t \quad \varepsilon_t \sim N(0, \sigma_t^2) \\ \sigma_t^2 &= \omega + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2.\end{aligned}$$

V rovnici podmienenej strednej hodnoty sú výnosy y_t modelované iba konštantou a nekorelovaným bielym šumom ε_t . Tento model zvyčajne postačuje na popísanie podmienenej strednej hodnoty výnosov, pretože mnohé finančné časové rady nepotrebujú rozšírenie, ktoré ARMA modely prinášajú. Podmienená volatilita je zase modelovaná konštantou a váženým priemerom minulej hodnoty volatility a štvorca rezíduí.

Kapitola 2

Neurónové siete

Výskum neurónových sietí bol podmienený hlbším poznaním fungovania ľudského mozgu. O mozgu by sa dalo povedať, že je to vlastne komplexný, nelineárny a paralelný počítač. Svoje základné stavebné jednotky má usporiadané veľmi dômyselným spôsobom, ktorý mu umožňuje vykonávať určité výpočty oveľa rýchlejšie ako ktorýkoľvek počítač. Práve táto skutočnosť motivovala vedcov ku vytvoreniu umelých neurónových sietí, ktoré by modelovali spôsob, akým ľudský mozog rieši určité úlohy.

Existuje viacero definícií, ktoré sa snažia presne vymedziť pojem neurónovej siete. Jedna z nich je táto:

Neurónová sieť je masívny paralelný procesor pozostávajúci z jednoduchých procesných jednotiek, ktorý má sklon k uchovávaniu experimentálnych poznatkov a ku ich sprístupňovaniu na ďalšie použitie. Podobá sa na ľudský mozog v dvoch ohľadoch:

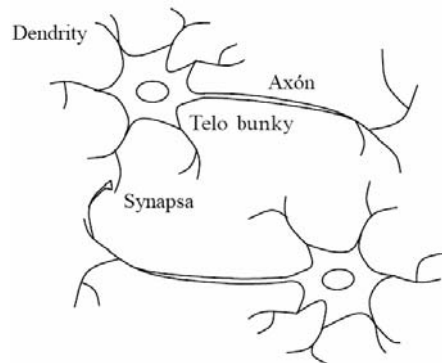
- 1. Poznatky sú získané neurónovou sieťou procesom učenia.*
- 2. Sila medzineurónových spojení vyjadrená veľkosťou synaptických váh slúži na ukladanie znalostí. [8]*

Pri neurónových sieťach zohráva veľmi dôležitú úlohu paralelizmus. V súčasnosti už dokážeme vytvoriť vysoko výkonné neuróny, ktoré prekonávajú schopnosti ľudských neurónov. Sila ľudského mozgu spočíva v obrovskom počte neurónov a medzineurónových spojení. Podľa výskumov má ľudský mozog asi 10^{12} neurónov, pričom počet medzineurónových spojení je približne 10^4 na každý neurón. Takúto masívnu paralelnú neurónovú sieť ešte nedokážeme na dnešných počítačoch nasimulovať.

2.1. Neurón

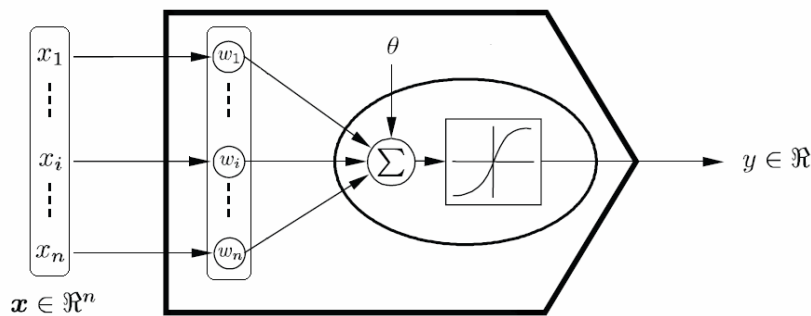
Neurón predstavuje základnú stavebnú jednotku neurónových sietí. Keďže umelými neurónovými sieťami simulujeme ľudský mozog, budeme umelý neurón ilustrovať na biologickom neuróne (obrázok 1). Nervová bunka sa skladá z tela a niekoľkých výbežkov. V tele bunky prebiehajú zložité biochemické deje, ktoré zabezpečujú spracovanie signálov. Výbežky neurónu možno rozdeliť na dva typy: dendrity a jeden axón. Dendrity predstavujú z informatického hľadiska vstupnú časť, pretože sa na ne prenáša signál z iných buniek. Axón tvorí naopak výstupný kanál z neurónu, ktorým sa šíri signál k iným nervovým bunkám.

Spojenie medzi dvoma neurónmi sa nazýva synapsa. Slúži nielen na prenos signálu z axónu jednej bunky na dendrity inej bunky, ale prenášané signály sa tu môžu zosilniť alebo zoslabiť.



Obrázok 1: Zjednodušený model biologického neurónu

Diagram na obrázku 2 predstavuje model neurónu, ktorý tvorí základ umelých neurónových sietí. Každý signál x_j vstupujúci do neurónu je prenášobný synaptickou váhou w_j . Tie na rozdiel od synapsí v ľudskom mozgu môžu nadobúdať kladné aj záporné hodnoty. Ďalšou dôležitou súčasťou je Σ , ktorá predstavuje lineárnu kombináciu vstupných signálov vyvážených zodpovedajúcimi synaptickými váhami. Tvorí vstup do aktivačnej funkcie, ktorá ohraňuje hodnoty výstupu z neurónu.



Obrázok 2: Zjednodušený matematický model neurónu

Model neurónu môže obsahovať aj prah neurónu, označený θ . V závislosti od znamienka môže znížiť alebo zvýšiť hodnotu vstupujúcu do aktivačnej funkcie. Matematicky môžeme neurón popísať nasledovnými rovnicami

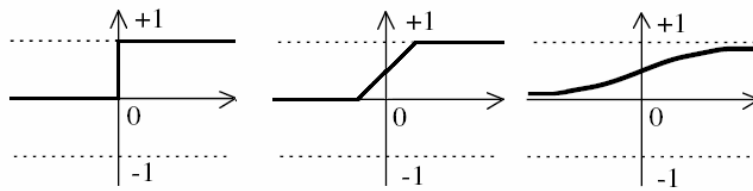
$$net = \sum_{j=1}^n w_j x_j + \theta$$

$$y = \varphi(net),$$

kde (x_1, \dots, x_n) je vstupný signál, (w_1, \dots, w_n) sú synaptické váhy, net je afinná kombinácia vstupných signálov a prahu neurónu θ , φ je aktivačná funkcia a y je výstup z neurónu. Práh neurónu predstavuje špeciálny parameter, ktorý nepochádza z iných neurónov. Môžeme ho však chápať ako ďalší vstup x_0 s hodnotou 1 a váhou w_0 rovnou θ .

Aktivačné funkcie

I keď existuje veľa aktivačných funkcií, rozlišujeme tri základné typy aktivačných funkcií: prahové, po častiach lineárne a sigmoidálne funkcie. Sú zobrazené na obrázku 3.



Obrázok 3: Zľava prahová funkcia, po častiach lineárna funkcia, sigmoidálna funkcia

Výber aktivačnej funkcie je podmienený úlohou, ktorú neurónové siete riešia. K najčastejšie používanému druhu aktivačných funkcií patria sigmoidálne funkcie. Ide o ostro rastúce funkcie, ktoré predstavujú akúsi rovnováhu medzi lineárnym a nelineárnym správaním. Príkladom sigmoidálnej funkcie je logistická funkcia

$$\varphi(\text{net}) = \frac{1}{1 + e^{-a \text{net}}}$$

kde a predstavuje sklon funkcie. Hlavnou výhodou týchto funkcií je vlastnosť spojitosti a diferencovateľnosti, ktorá je v teórii neurónových sietí veľmi dôležitá. Navyše sa ich derivácia dá vyjadriť v jednoduchom tvare ako

$$\varphi'(\text{net}) = a \varphi(\text{net})(1 - \varphi(\text{net})).$$

Tieto funkcie majú obor hodnôt $[0,1]$. Niekedy však potrebujeme aktivačné funkcie s rozsahom hodnôt $[-1,1]$. To sa dá ľahko dosiahnuť transformáciou predošlých funkcií. V empirickej časti budeme využívať sigmoidálnu aktivačnú funkciu transformovanú na interval $[-1,1]$. Konkrétne pôjde o funkciu v tvare

$$\varphi(v) = \frac{2}{1 + e^{-2v}} - 1.$$

2.2. Neurónové siete

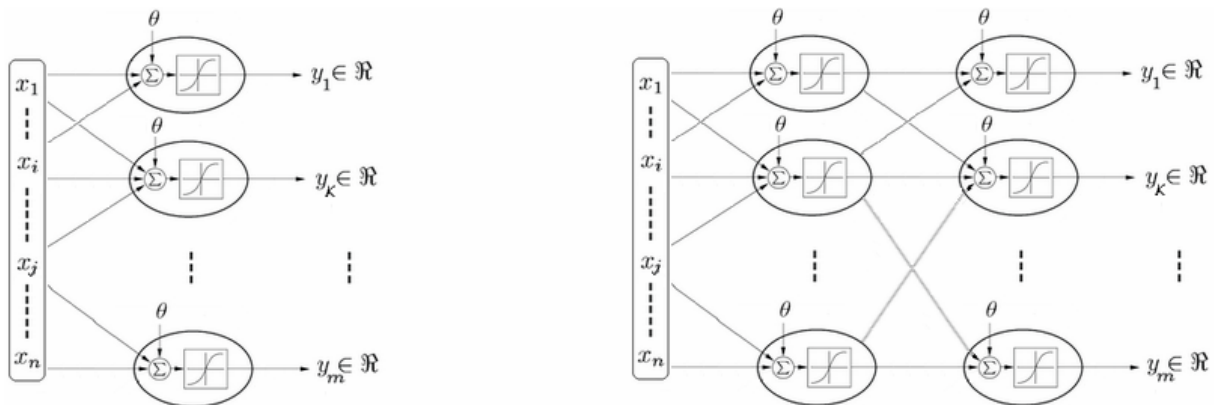
Možnosti využitia samotného neurónu sú značne limitované. Rovnako ako ľudský mozog i umelé neurónové siete získavajú svoje schopnosti zo zložitej paralelnej štruktúry. Spôsob, akým sú neuróny usporiadané do neurónovej siete je úzko spojený s učiacim algoritmom používaným na tréning siete. Vo všeobecnosti rozlišujeme dve triedy architektúry neurónových sietí.

Dopredné neurónové siete

Táto architektúra neurónových sietí sa vyznačuje tým, že obsahuje iba dopredné spojenia medzi neurónmi, t.j. signál sa šíri po orientovaných synaptických prepojeniach len jedným smerom a to dopredu. Každý neurón teda vysiela signály iba na neuróny v nasledujúcej vrstve. Spojenia do predchádzajúcej vrstvy ani v rámci jednej vrstvy neexistujú.

Na obrázku 4 je znázornená jednovrstvová dopredná neurónová sieť. Ide o najjednoduchšiu formu neurónovej siete, ktorá sa skladá z vrstvy vstupných neurónov

a vrstvy výstupných neurónov. Externé vstupy sú dodávané do siete vstupnými neurónmi. Tieto neuróny však nevykonávajú žiadne matematické operácie, preto ich neuvažujeme ako samostatnú vrstvu. Výstupná vrstva zodpovedá konečnému výstupu z neurónovej siete. Okrem jednovrstvových existujú aj viacvrstvové neurónové siete. Táto architektúra neurónových sietí sa od tej predchádzajúcej líši v tom, že obsahuje aj skryté vrstvy neurónov. Tie ju robia výkonnejšou v porovnaní s jednovrstvovou sieťou. Vďaka dodatočným synaptickým spojeniam dokáže sieť extrahovať z dát ďalšie informácie. Každá vrstva má pritom svoje vlastné váhy, prahy neurónov a aktivačné funkcie. Signál sa šíri sieťou od vstupnej vrstvy cez skrytú až po výstupnú vrstvu. Výstupný signál neurónov výstupnej vrstvy predstavuje celkovú odpoveď siete na vstupy zadávané do vstupnej vrstvy.



Obrázok 4: Zľava jednovrstvová a viacvrstvová dopredná neurónová sieť

Rekurentné neurónové siete

Neurónové siete spomínané doposiaľ sú tzv. dopredné neurónové siete, pretože neurónové spojenia smerujú len jedným smerom a to dopredu. V prípade rekurentných neurónových sietí máme nielen dopredné spojenia ale i spätnú väzbu. Tá dodáva tomuto typu sietí vlastnosti, ktoré nie sú obsiahnuté v dopredných neurónových sieťach, konkrétne vlastnú pamäť. Sú schopné ukladať informácie na ďalšie použitie.

2.3. Neurónové siete vyššieho rádu

Doposiaľ spomínané typy neurónových sietí dokážu zachytiť iba korelácie prvého rádu medzi jednotlivými elementmi vstupov. V mnohých prípadoch by sme však potrebovali nástroj schopný zachytiť aj korelácie vyššieho rádu. Za týmto účelom boli vytvorené neurónové siete vyššieho rádu (skrátene HONN z angl. *higher order neural networks*). Na predpovedanie výmenných kurzov budeme používať neurónové siete druhého rádu, ktoré patria ku najčastejšie používanému typu HONN. Korelácie a vzťahy medzi jednotlivými elementmi vstupov identifikujú ich vzájomným prenásobením. Táto operácia môže byť vykonaná ešte pred vstupom dát do siete alebo vhodnou voľbou neurónov v skrytej vrstve siete. Majme vektor vstupných dát x_1, \dots, x_n . V prvom prípade zoberieme za vstup siete nielen túto n -ticu, ale aj všetky $x_i x_j$ pre $i \neq j$. Matematický popis výstupného neurónu v jednovrstvovej sieti bude vyzeráť nasledovne

$$y = \varphi \left(w_0 + \sum_{i1}^n w_{i1} x_{i1} + \sum_{i1,i2}^n w_{i1,i2} x_{i1} x_{i2} \right),$$

kde $x = [x_1, x_2, \dots, x_n]^T$ je vektor vstupov, y je výstup zo siete a $\varphi(\cdot)$ je aktivačná funkcia. V druhom prípade ide o použitie špeciálnych neurónov. Neurónové siete prvého rádu, teda klasické neurónové siete, sú vytvárané lineárnymi neurónmi, ktoré vytvárajú afinnú kombináciu vstupných dát a váhového vektora. HONN majú podobnú architektúru ako dopredné neurónové siete, pričom lineárne neuróny sú nahradené násobnými neurónmi.

Počet neurónov v HONN je nižší ako pri klasických neurónových sieťach, čo znižuje počet trébovaných váh. Na druhej strane počet váh rastie exponenciálne s počtom vstupov. Preto sa architektúry vyššieho ako tretieho rádu používajú iba veľmi zriedka.

2.4. Učenie

Veľmi dôležitou vlastnosťou neurónových sietí je schopnosť učiť sa a zlepšovať svoj výkon učením. Pojem učenia neurónovej siete je synonymom pojmu adaptácie. Uskutočňuje sa adaptáciou váh a prahov neurónov, pričom znalosti získané učením sa ukladajú do synaptických váh neurónovej siete. Adaptácia parametrov siete prebieha na základe istých pravidiel, ktoré determinujú typ učenia. Tieto pravidlá vytvárajú učiaci algoritmus neurónovej siete. V zásade sa jednotlivé učiace algoritmy odlišujú v spôsobe, ktorým je formulovaná zmena váh. Vo všeobecnosti rozlišujeme dva typy učenia

- *kontrolované učenie*

Spôsob zmeny synaptických váh je ovplyvnený prítomnosťou učiteľa v procese učenia. V praxi to znamená, že do neurónovej siete musíme zadať nielen vstup ale i k vstupu prislúchajúci výstup. Do tejto skupiny zaraďujeme viacero prístupov ku zmene synaptických váh. My sa v tejto práci budeme zaoberať učením na základe korekcie chyby.

- *nekontrolované učenie*

Pri tomto type učenia ponúkžeme neurónovej sieti iba vstup. Sieť ho spracuje a sama určí výstup. Proces učenia pokračuje dovtedy, kým sa sieť nedostane do stabilného stavu, teda kým zmena synaptických váh nebude dostatočne malá.

Kapitola 3

Učenie na základe korekcie chyby

Ide o metódu, ktorá porovnáva výstup siete s požadovaným výstupom. Existuje viacero spôsobov, ktorými môžeme definovať výšku chybového signálu. Najčastejšie je však jeho výška určená rozdielom požadovanej a vypočítanej hodnoty. Práve na základe veľkosti chybového signálu sa spustí mechanizmus krokov, ktoré upravujú synaptické váhy tak, aby priblížili vypočítaný stav siete ku požadovanej hodnote. Najčastejšie používaným učiacim algoritmom aplikovaným pri učení na základe korekcie chyby je metóda spätného šírenia chyby.

3.1. Metódy riešenia úloh na voľný extrém

Učiace algoritmy sú založené na minimalizácii chybového signálu. Preto kým prejdeme na samotnú metódu spätného šírenia chyby, povieme si niečo o metódach používaných na riešenie úloh na voľný extrém. Predpokladajme, že máme diferencovateľnú chybovú funkciu $E(w)$, ktorá je funkciou neznámeho parametra w . Riešime optimalizačný problém, ktorého cieľom je minimalizovať chybovú funkciu $E(w)$ vzhľadom na váhový vektor w , teda

$$\text{Min} \left\{ E(w) \mid w \in \mathfrak{R}^n \right\}.$$

Jedná sa o úlohu na voľný extrém. Takmer všetky metódy riešenia takéhoto typu úloh sú iteračné, teda generujú postupnosť bodov konvergujúcu k optimálnemu riešeniu úlohy respektíve aspoň k bodu lokálneho minima. Dôležitou vlastnosťou iteračných metód je ich schopnosť generovať v každom kroku lepšiu aproximáciu, ako bola tá predošlá. Metódy, ktoré znižujú hodnotu účelovej funkcie každou iteráciou sa nazývajú spádové.

Hľadané minimum w^* , musí spĺňať nutnú podmienku optimality, tj.

$$\nabla E(w^*) = 0,$$

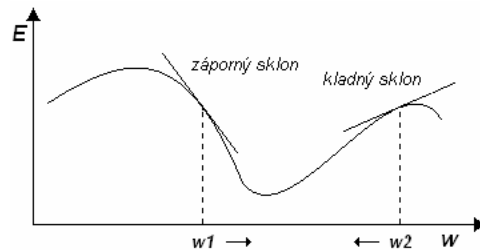
kde

$$\nabla E = \left(\frac{\partial E}{\partial w_1} \quad \frac{\partial E}{\partial w_2} \quad \dots \quad \frac{\partial E}{\partial w_m} \right)^T.$$

V ďalších úvahách budeme gradient označovať symbolom g , teda $g = \nabla E$.

3.1.1. Gradientová metóda najväčšieho spádu

Patrí medzi základné metódy na minimalizáciu funkcií. S využitím gradientu účelovej funkcie konštruuje postupnosť bodov, ktorá za istých predpokladov dokonverguje do stacionárneho bodu. Smer najväčšieho poklesu chybovej funkcie je určený záporným gradientom.



Obrázok 5: Ilustrácia gradientovej metódy najväčšieho spádu

Obrázok 5 schematicky zobrazuje vzťah medzi chybovou funkciou a váhovým vektorom. Ak je hodnota gradientu záporná, na zníženie chybovej funkcie musíme zvýšiť váhy. Naopak, ak je hodnota gradientu kladná, znížením váh sa priblížime k hodnote minima. Matematický zápis metódy bude mať tvar

$$w(t+1) = w(t) - \eta g(t),$$

kde η predstavuje mieru učenia. Zvyčajne je $\eta \in \mathfrak{R}$, $0 < \eta \leq 0.5$.

Metóda najväčšieho spádu konverguje iba lineárne, čo ju robí neefektívnou na riešenie praktických úloh. I keď skoro vždy dokonverguje do minima, potrebuje na to príliš veľký počet iterácií. Ďalším nedostatkom je konštantná dĺžka kroku. V nasledujúcej časti popíšeme sofistikovanejšie metódy, ktoré konvergujú rýchlejšie.

3.1.2. Newtonova metóda

Jej cieľom je minimalizovať chybovú funkciu $E(w)$ okolo približného riešenia $w(t)$. Keďže metóda využíva Taylorov rozvoj do druhého rádu, musíme uvažovať iba chybové funkcie so spojitými druhými parciálnymi deriváciami. Príslušný Taylorov rozvoj má tvar

$$E(w) \cong E(w(t)) + g^T(t)(w - w(t)) + \frac{1}{2}(w - w(t))^T H(t)(w - w(t)),$$

kde vektorom g označujeme gradient chybovej funkcie $E(w)$, maticou H zase Hessovu maticu, teda $H = \nabla^2 E$. V Newtonovej metóde je nové približné riešenie $w(t+1)$ definované ako minimum tejto kvadratickej funkcie. Keďže musíme predpokladať kladnú definitnosť Hessovej matice, obmedzíme sa iba na konvexné chybové funkcie. Newtonova metóda má potom tvar

$$w(t+1) = w(t) - H^{-1}(t)g(t).$$

Metóda za istých rozumných predpokladov konverguje kvadraticky ku bodu lokálneho minima. Problém môže nastať, ak iteračné schéma začne príliš ďaleko od hľadaného minima. Vtedy môže vykazovať nevyspytateľné správanie. Ďalším nedostatkom Newtonovej metódy je nutnosť výpočtu Hessovej matice druhých parciálnych derivácií. Existuje však trieda tzv.

kvázinewtonovských metód, ktorá namiesto výpočtu inverznej Hessovej matice používa iba jej postupnú aproximáciu. Patrí sem BFGS (Broyden, Fletcher, Goldfarb a Shanno) metóda, pri ktorej určíme postupnú aproximáciu vzorcom

$$H_+ = H + \frac{1}{p^T y} y y^T - \frac{1}{p^T H p} H^T p^T p H,$$

respektíve

$$H_+^{-1} = H^{-1} + \left(1 + \frac{y^T H^{-1} y}{p^T y}\right) \frac{1}{p^T y} p p^T - \frac{1}{p^T y} (H^{-1} y p^T + p y^T H^{-1}),$$

kde

$$p = w_+ - w$$

$$y = g_+ - g.$$

3.1.3. Levenberg-Marquardtova metóda

Každá z vyššie spomenutých metód má svoje výhody. Gradientová metóda zaručene skonverguje k minimu aj zo vzdialeného počiatočného bodu, zatiaľ čo Newtonova metóda pre dostatočne blízky počiatočný bod konverguje kvadraticky. Levenberg prišiel s nápadom spojiť tieto dva prístupy v jednej metóde, ktorá by mala obe tieto výhody. Označme λ ako parameter vyjadrujúci silu spojenia týchto dvoch metód. Iteračná schéma Levenbergovej metódy má tvar

$$w(t+1) = w(t) - (H(t) + \lambda I)^{-1} g(t),$$

kde I je jednotková matica. Algoritmus Levenbergovej metódy upravuje hodnotu λ v závislosti od výšky chybovej funkcie. V prípade nárastu chyby nefunguje dobre kvadratická aproximácia a my sme ďaleko od minima. Takže zvýšime hodnotu parametra λ , čím dáme priestor metóde najväčšieho spádu. Naopak, ak chyba klesá, aproximácia funguje správne a my sa pravdepodobne dostávame bližšie ku hodnote minima. Preto znížime hodnotu parametra λ a dáme väčší priestor kvadratickej aproximácií.

Ako sme už spomenuli, pre veľké hodnoty λ aplikujeme vlastne metódu najväčšieho spádu. Marquardtov prínos spočíval v tom, že i v tejto situácii by sme mohli ťažiť z Hessovej matice, ktorú sme odhadli. V iteračnej schéme nahradil jednotkovú maticu I diagonálovú Hessiánu, čím dostal rovnicu

$$w(t+1) = w(t) - (H(t) + \lambda \text{diag}(H(t)))^{-1} g(t).$$

Tu sa tiež stretávame s podobným problémom výpočtovej náročnosti ako pri Newtonovej metóde. Preto bola i Levenberg-Marquardtova metóda upravená tak, aby nebolo nutné počítať Hessovu maticu druhých parciálnych derivácií. Aproximácie použité v tomto prípade vychádzajú z Gauss-Newtonovej metódy. Tá sa používa na hľadanie minima funkcií, ktoré sú súčtom štvorcov nelineárnych funkcií. Pre tento typ funkcií môžeme použiť aproximácie

$$H = J^T J \quad \text{a} \quad g = J^T e,$$

kde e označuje chybový vektor a J zase Jakobián obsahujúci prvé derivácie chybového vektora podľa jednotlivých váh. Keďže výpočet Jakobiánu je jednoduchší ako výpočet Hessovej matice, predstavuje táto aproximácia zníženie výpočtovej náročnosti metódy.

S využitím tejto aproximácie bude mať iteračná schéma Levenberg-Marquardtovej metódy tvar

$$w(t+1) = w(t) - (J(t)^T J(t) + \lambda I)^{-1} J(t)^T e(t),$$

resp.

$$w(t+1) = w(t) - (J(t)^T J(t) + \lambda \text{diag}(J(t)^T J(t)))^{-1} J(t)^T e(t).$$

Je dôležité zdôrazniť, že ide o heuristickú metódu. Nie je optimálna v zmysle rýchlosti konvergencie ani výslednej chyby [24]. Je to proste dobre vymyslená optimalizačná procedúra. Napriek tomu dosahuje v praxi veľmi slušné výsledky. Stala sa akousi štandardnou metódou na optimalizáciu stredne veľkých neurónových sietí. Avšak pre väčšie systémy sa stáva výpočtovo náročnou.

3.2. Delta pravidlo

Ide o učiace pravidlo používané pri úprave synaptických váh jednovrstvových neurónových sietí. Uvažujme, že máme sieť pozostávajúcu z m vstupných a jedného výstupného neurónu k s lineárnou aktivačnou funkciou. Neurón k prijíma signál zo vstupnej vrstvy vo forme vektora x . Výstupný signál y_k je porovnaný s požadovaným výstupom d_k , pričom výšku chybového signálu určíme ako rozdiel týchto dvoch veličín. Chybová funkcia, ktorú budeme uvažovať aj pri ďalších odvozeniach, bude mať tvar

$$E = \frac{1}{2} e_k^2 = \frac{1}{2} (d_k - y_k)^2.$$

Minimalizáciou chybovej funkcie gradientovou metódou najväčšieho spádu dostaneme učiace pravidlo známe pod názvom delta pravidlo. Odvodenie pravidla vychádza z toho, že zmena váhy w_{kj} je nepriamo úmerná parciálnej derivácii chybovej funkcie podľa w_{kj}

$$\Delta w_{kj} = -\eta \frac{\partial E}{\partial w_{kj}}.$$

Úpravami dostaneme nasledovný vzorec pre zmenu váh

$$\Delta w_{kj} = \eta e_k x_j.$$

Postup odvodenia tohto vzorca je totožný s postupom odvodenia metódy spätného šírenia chyby pre výstupné neuróny, ktorý je uvedený v prílohe B. Delta pravidlo predstavuje základný kameň učenia neurónu, ktoré bude bežať dovtedy, kým systém nedosiahne stabilný stav. V tom bode dôjde k prerušeniu procesu učenia.

3.3. Metóda spätného šírenia chyby

V predošlej časti zadefinované delta pravidlo predstavuje základ učenia metódou spätného šírenia chyby. Cieľom učenia je dosiahnuť taký stav, v ktorom sa skutočný výstup siete rovná požadovanému výstupu pre všetky vzory z tréningovej množiny, respektíve, v ktorom sa mu

približuje čo najviac. Štandardná metóda je odvodená na chybovej funkcii definovanej ako súčet štvorcov chýb jednotlivých výstupných neurónov, teda

$$E = \frac{1}{2} \sum_{k=1}^N (d_k - y_k)^2.$$

Metóda budeme minimalizovať chybovú funkciu vzhľadom na veľkosť synaptických váh a prahových koeficientov jednotlivých neurónov. Štandardná metóda spätného šírenia chyby je algoritmus využívajúci gradientovú metódu najväčšieho spádu, v ktorom sú váhy modifikované v smere záporného gradientu chybovej funkcie. Názov spätné šírenie chyby sa vzťahuje ku spôsobu, akým je vypočítaná veľkosť gradientu. Celý postup odvedenia tejto metódy je popísaný v prílohe B. Výsledná formula má tvar

$$w_{kj}^l(t+1) = w_{kj}^l(t) + \eta \delta_k^l y_j^{l-1},$$

$$\delta_k^l = \begin{cases} e_k^L \varphi_k'(net_k^L) & \text{pre neurón } k \text{ vo výstupnej vrstve } L \\ \varphi_k'(net_k^L) \sum_i \delta_i^{l+1} w_{ik}^{l+1} & \text{pre neurón } k \text{ v skrytej vrstve } l \end{cases}$$

kde δ_k^l je zovšeobecnený chybový signál, ktorý sa spätne šíri na všetky váhy prichádzajúce ku k -temu neurónu.

Metóda spätného šírenia chyby sa skladá z dvoch výpočtových fáz. V prvej fáze postupujeme od vstupnej vrstvy, cez skryté vrstvy až po výstupnú vrstvu. Pri prechode sieťou sa v jednotlivých vrstvách vypočítajú stavy každého z neurónov. Výstup siete sa porovná s požadovanou hodnotou výstupu, čím získame veľkosť chybového signálu. Druhá fáza začína vo výstupnej vrstve, odkiaľ spätne šíri zovšeobecnený chybový signál vrstvu po vrstve. Pre výstupnú vrstvu je δ spočítané priamo. Pomocou tejto vypočítanej hodnoty určíme veľkosť δ pre predchádzajúcu skrytú vrstvu. Takýmto rekurzívnym spôsobom vypočítame všetky zovšeobecnené chybové signály, z ktorých potom určíme zmenu váh v jednotlivých vrstvách siete.

V rovnici pre úpravu synaptických váh vystupuje parameter η , ktorý predstavuje mieru učenia. Čím je hodnota η menšia, tým menšie budú zmeny váh, čo sa zase odzrkadlí v hladšej trajektórii vo váhovom priestore, ale iba za cenu spomalenia konvergence neurónovej siete. Naopak veľké hodnoty parametra η môže viesť ku osciláciám. Pri odvodení metódy spätného šírenia chyby sme mieru učenia uvažovali ako konštantnú veličinu. V praxi sa však jej hodnota zvykne meniť. Jednoduchou metódou, ako zvýšiť mieru učenia a znížiť riziko oscilácií je modifikovanie delta pravidla do tvaru

$$\Delta w_{kj}(t) = \Delta w_{kj}(t) + \alpha \Delta w_{kj}(t-1),$$

kde α je kladný parameter nazývaný momentum. Jeho veľkosť sa pohybuje rádovo od 10^{-2} až 10^{-1} . Vďaka momentu sa pri minimalizácii môžeme vyhnúť lokálnym minimám.

Majme trérovaciu množinu $\{(x^1, d^1)(x^2, d^2) \dots (x^P, d^P)\}$. Potom algoritmus trérovania doprednej neurónovej siete pomocou spätného šírenia chýb môže mať nasledovný tvar:

- 1) Zvolíme $\alpha \in (0, 1)$. Počiatočné váhy vrátane prahov neurónov inicializujeme ako nízke náhodné čísla. Poradové číslo epochy, t.j. prechodu cez trérovaciu množinu, označíme ako h , index vzoru ako p , E_{avg} ako priemernú chybu

$$E_{avg} = \frac{1}{P} \sum_{p=1}^P E_p,$$

kde E_p je výška chybovej funkcie pre vzor p , t.j. $E_p = \frac{1}{2} \sum_{k=1}^N (d_k^p - y_k^p)^2$. Počítadlá a chybu nastavíme takto: $h = 1, p = 1, E = 0$. Ešte zvolíme malé kladné číslo ε , ktoré nám bude slúžiť na zastavenie učenia.

- 2) Za vstup zoberieme vzor x^p a vypočítame výstup y^p .
- 3) Pre všetky neuróny v skrytej a výstupnej vrstve vypočítame δ_k
- 4) Modifikujeme váhy neurónov $w_{kj} \leftarrow w_{kj} + \eta \delta_k y_j$
- 5) Ak $p < P$, položíme $p = p + 1$ a prejdeme na krok 2. Inak prejdeme na krok 6.
- 6) Po prechode celou trérovacou množinou vypočítame priemernú chybu E_{avg} jednej epochy učenia. Ak je $E_{avg} < \varepsilon$, ukončíme učenie. Inak náhodne premiešame vzory v trérovacej množine, aby v každej epoche učenia vstupovali do siete v inom poradí. Položíme $E = 0, p = 1, h = h + 1$ a prejdeme na krok 2. [17]

Tento algoritmus zodpovedá sekvenčnému učeniu, pri ktorom upravujeme váhy po zadaní každého trérovacieho vzoru. Okrem toho existuje ešte jeden typ, tzv. učenie v dávkach, pri ktorom meníme váhy až po prechode celej trérovacej množiny sieťou. Chybu jednej epochy definujeme ako priemernú chybu dosiahnutú pri jednotlivých vzoroch, tj.

$$E_{avg} = \frac{1}{2P} \sum_{p=1}^P \sum_{k=1}^N (d_k^p - y_k^p)^2.$$

Pri zmene váh vychádzame z gradientu E_{avg} , teda

$$\Delta w_{kj} = -\eta \frac{\partial E_{avg}}{\partial w_{kj}} = -\frac{\eta}{P} \sum_{p=1}^P e_k^p \frac{\partial e_k^p}{\partial w_{kj}}.$$

Pri výpočte derivácie $\partial e_k^p / \partial w_{kj}$ postupujeme rovnakým spôsobom ako pri sekvenčnom učení. Detaily sú znova v prílohe B.

Kapitola 4

Meranie výkonnosti

Najbežnejším kritériom na vyhodnotenie predikčnej schopnosti siete je stredná kvadratická chyba (skrátene MSE z angl. *Mean Squared Error*). Jej veľkosť dostaneme vzorcom. Porovnáva, do akej miery sa líši výstup siete od požadovaného výstupu. Čiže ľubovoľná predikcia, ktorej hodnoty sú blízko skutočným hodnotám, nadobúda malé hodnoty strednej kvadratickej chyby.

MSE nám však poskytuje iba akési kvantitatívne vyjadrenie presnosti predpovede. Z ekonomického hľadiska však existuje oveľa dôležitejšie kritérium a tým je dosiahnutie zisku z investičných rozhodnutí vyplývajúcich z predpovede. Pri investičných rozhodnutiach je dôležité správne predpovedanie pohybov ceny aktíva, teda jeho nárastu resp. poklesu. Toto kritérium môžeme v prípade cien aktív vyjadriť štatistikou D_{stat} , ktorá meria smery a zlomové body v dátach. V prípade výnosov bude ekvivalentná štatistika S_{stat} vyjadrovať znamienkovú rovnosť medzi skutočnými a predpovedanými hodnotami. Majú nasledovný tvar

$$S_{stat} = \frac{1}{N} \sum_{t=1}^N a_t \times 100\%, \quad D_{stat} = \frac{1}{N} \sum_{t=1}^N b_t \times 100\%,$$

kde

$$a_t = \begin{cases} 1 & \text{ak } y_t \hat{y}_t > 0, y_t = \hat{y}_t = 0 \\ 0 & \text{inak} \end{cases}$$

kde

$$b_t = \begin{cases} 1 & \text{ak } (y_{t+1} - y_t)(\hat{y}_{t+1} - \hat{y}_t) \geq 0 \\ 0 & \text{inak} \end{cases}$$

Predstavujú prvý odhad výnosnosti. Nás však bude zaujímať najmä zisk dosiahnutý na základe predpovedaných hodnôt. Smerujeme ku analýze ziskov a strát, ktorá simuluje reálny tok peňazí z investície na finančných trhoch. Preto ďalším kritériom, ktoré prichádza do úvahy, je ročná miera výnosu. Bez uvažovania nákladov je ročná výnosová miera určená vzťahom

$$R = \left(\left(\frac{M}{S} \right)^{12/N} - 1 \right) \times 100\%,$$

kde R je ročná výnosová miera, M je množstvo peňazí získaných v posledný deň testovacieho obdobia, S je množstvo peňazí použitých na obchodovanie v prvý deň a N je dĺžka testovacieho obdobia v mesiacoch. Výpočet R je založený na voľbe konkrétnej obchodnej stratégie. Vo všeobecnosti však pri výpočte vychádzame z istej sumy peňazí, ktorú využívame na nákup resp. predaj v závislosti od konkrétnej obchodnej stratégie a predpovedí neurónovej

siete. Na základe sumy, ktorú budeme mať na konci sledovaného obdobia, vypočítame ročnú výnosovú mieru.

Obchodné stratégie

Ako sme už spomenuli, ročná výnosová miera závisí od konkrétnej stratégie investora. Keďže neexistuje žiadna stopercentná metóda predpovedania, maximálny výnos dosiahneme iba dobrou obchodnou stratégiou, ktorá z predpovede vyťaží maximum. V tejto práci budeme využívať dve obchodné stratégie. Prvá vychádza z rozdielu medzi predpoveďami, zatiaľ čo druhá berie rozdiel medzi predpovedanou a aktuálnou hodnotou.

Stratégia 1

ak $(\hat{y}_{t+1} - \hat{y}_t) > 0$ vtedy *kúpiť* inak *predať*

Stratégia 2

ak $(\hat{y}_{t+1} - y_t) > 0$ vtedy *kúpiť* inak *predať*

Dôvod, pre ktorý dávajú tieto dve stratégie rôzne výsledky, spočíva v šume. Ak označíme šum v čase t ako δ_t , potom platí

$$y_t = \hat{y}_t + \delta_t.$$

To, ktorá z týchto stratégií je lepšia, závisí od $|\delta_{t+1} - \delta_t|$ a $|\delta_{t+1}|$. Ak majú δ_{t+1} a δ_t rozdielne znamienka, vtedy je lepšia stratégia 1. V opačnom prípade je lepšia stratégia 2. [31]

Ak by sme sa chceli čo najviac priblížiť realite, bolo by potrebné zahrnúť do výpočtu ročnej výnosovej miery aj náklady. Nás budú zaujímať obe ročné výnosové miery, teda bez nákladov i s nákladmi na úrovni 0.25% z každej transakcie.

Štandardy

Kľúčovým momentom predpovedania vo financiách je vyhodnotenie predikčnej schopnosti algoritmu. V predchádzajúcej časti sme si popísali viaceré miery výkonnosti neurónovej siete, ktoré však samé o sebe nie sú dostatočne uspokojivým zhodnotením algoritmu. Na to je potrebné stanoviť si istý štandard, ktorým je zvyčajne iný predikčný algoritmus použitý ako porovnanie. V nasledujúcej časti si ich zopár predstavíme.

Akýkoľvek úspešný predikčný algoritmus by mal prekonať naivnú predikciu

$$\hat{y}_{t+1} = y_t,$$

kde y_t je aktuálna hodnota a \hat{y}_{t+1} je predpovedaná hodnota. Za najlepšiu predpoveď budúcej hodnoty považuje súčasný stav, čo je v súlade s Hypotézou efektívneho trhu. V prípade modelovania finančných aktív môžeme porovnať výnosy získané na základe predpovede s inými investičnými možnosťami. My budeme uvažovať jednoduchú stratégiu Buy & Hold, kedy na začiatku testovacieho obdobia nakúpime za počiatočný objem peňazí istý objem aktív a potom ich na konci tohto obdobia predáme. Bude nás zaujímať nielen výška zisku ale aj zodpovedajúca ročná výnosová miera.

Kapitola 5

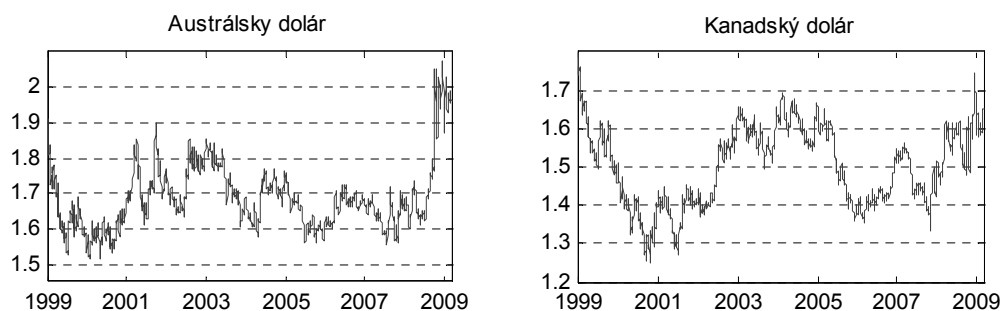
Dáta

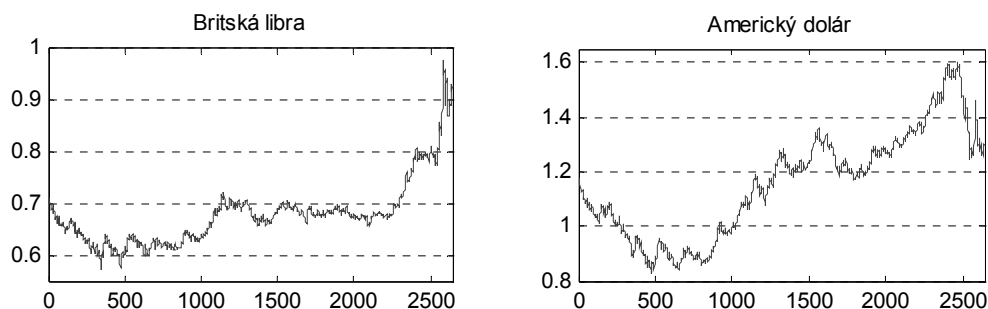
Pri technickej analýze finančných časových radov sa na modelovanie využívajú historické hodnoty daného radu. Za vstupy sa berú nielen ceny a objemy obchodovaných aktív, ale často aj údaje, ktoré vznikli ich transformáciou. Do tejto skupiny patria výnosy a volatilita ale i ďalšie technické indikátory. Hlavne výnosy sa pri modelovaní používajú pomerne často, keďže predstavujú spôsob, ako transformovať pôvodne nestacionárne dáta na stacionárne. My budeme v tejto práci pracovať s logaritmickým výnosom, ktorý definujeme ako

$$R(t) = \log \frac{y(t)}{y(t-1)}.$$

V prípade finančných dát sa pomerne často stretávame s ich nízkou kvalitou, hlavne s chýbajúcimi dátami. Bežným spôsobom ako prekonať tieto prekážky, je nechať sieť akceptovať tieto chýbajúce hodnoty. Tá zvyčajne dokáže vyčítať z dát dôležitú dynamiku aj napriek týmto nedostatkom. My sme však zvolili iný prístup. Chýbajúce údaje sme nahradili hodnotami z predchádzajúceho časového okamihu, ktoré sme považovali za ich najlepšiu aproximáciu.

Na začiatku budeme vychádzať zo štyroch výmenných kurzov, konkrétne pôjde o výmenné kurzy eura voči austrálskemu doláru, kanadskému doláru, britskej libe a americkému doláru. Dáta berieme zo Štatistickej databázy Európskej Centrálnaj Banky (<http://sdw.ecb.europa.eu>). Sú za obdobie od 4.1.1999 do 17.3.2009, čo predstavuje časový rad za 2641 pracovných dní. Transformáciou denných výmenných kurzov na denné logaritmické výnosy zredukujeme ich počet na 2640. Budeme analyzovať vlastností týchto časových radov. Na základe výsledkov analýzy zvolíme dva výmenné kurzy, ktoré budeme predpovedať.





Obrázok 6: Grafické znázornenie vývoja výmenných kurzov

Vstupné dáta je potrebné ešte transformovať do intervalu $[0,1]$ prípadne $[-1,1]$, aby boli konzistentné s použitou aktivačnou funkciou. Existuje viacero metód škálovania, najčastejšie sa však stretávame s nasledovnou lineárnou transformáciou

$$y_{trans} = \varphi_{min} + \frac{y - y_{min}}{y_{max} - y_{min}} (\varphi_{max} - \varphi_{min}),$$

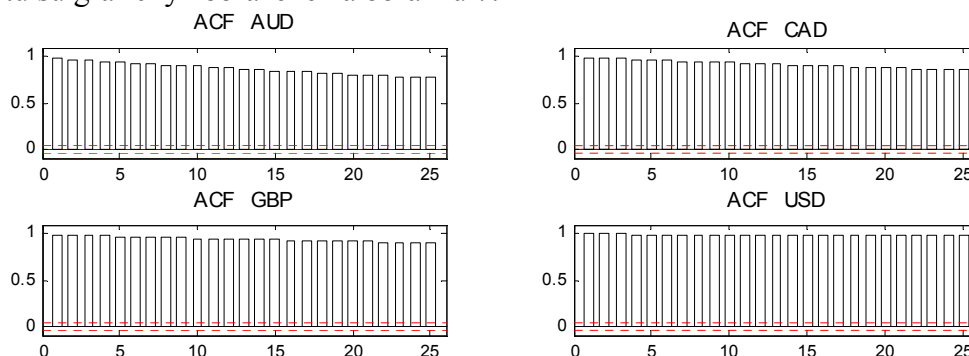
kde y_{trans} je transformovaná hodnota, φ_{min} resp. φ_{max} je minimum resp. maximum transferovej funkcie $\varphi(\cdot)$, y je pôvodná hodnota, y_{min} resp. y_{max} sú jej minimum resp. maximum.

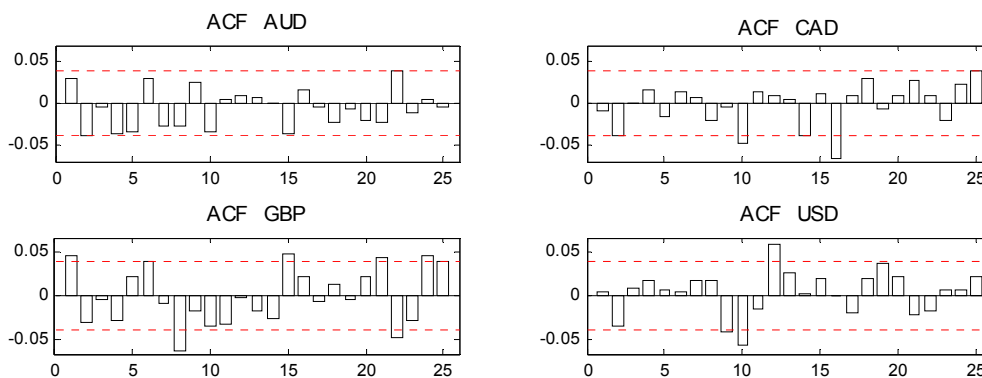
5.1. Štatistické vlastnosti dát

V tejto časti budeme analyzovať vlastnosti výmenných kurzov a ich výnosov. Základné štatistické vlastnosti výmenných kurzov dokážeme určiť z grafov znázornených na obrázku 6. Nás bude hlavne zaujímať prítomnosť pamäťových efektov v radoch vybraných výmenných kurzov, teda do akej miery je daná hodnota časového radu ovplyvnená predchádzajúcimi hodnotami. Tieto vzťahy budeme skúmať pomocou autokorelačnej funkcie a Hurstovho koeficientu.

Autokorelačná funkcia

O význame autokorelačnej funkcie sme sa v stručnosti zmieňovali v 1. kapitole. Autokorelačný koeficient časového radu pri omeškaní k je odhadnutý výberovým autokorelačným koeficientom. Určuje smer a stupeň lineárneho vzťahu medzi pozorovaným radom a radom posunutým o k hodnôt. Príslušné hodnoty výberového autokorelačného koeficientu sú graficky zobrazené na obrázku 7.





Obrázok 7: Autokorelogramy cien a výnosov výmenných kurzov

Z obrázku 7 vidíme, že samotné výmenné kurzy sú výrazne autokorelované, pričom autokorelácia je najsilnejšia v prípade výmenného kurzu EUR/USD a „najslabšia“ pre EUR/AUD. Vo všetkých štyroch prípadoch sú však efekty dlhodobej pamäte pomerne výrazné. To indikuje predpovedateľnosť výmenného kurzu. Z grafického znázornenia výmenných kurzov a z priebehu autokorelačných funkcií je navyše jasné, že ide o nestacionárne procesy. Nestacionarita týchto radov môže byť problémom pri klasických metódach predpovedania. Preto budeme na modelovanie využívať výnosy. Ich autokorelačné funkcie sú taktiež znázornené na obrázku 7. Naznačujú stacionárnosť, ktorú potvrdili aj výsledky rozšíreného Dickey-Fullerovho testu. Tie sú znázornené v tabuľke 1. Test vo všetkých štyroch prípadoch zamietá nulovú hypotézu, časové rady nemajú jednotkový koreň a teda sú stacionárne.

	AUD	CAD	GBP	USD
t-štatistika	-49.899163	-51.821128	-49.163071	-51.104808
p-hodnota	0.0001	0.0001	0.0000	0.0001

Tabuľka 1: Výsledky Dickey-Fullerovho testu stacionarity pre výnosy výmenných kurzov

Hurstov koeficient

Na odhad hodnoty Hurstovho koeficientu použijeme analýzu preškáľovaného rozsahu. Pri tejto metóde treba rozdeliť časový rad do navzájom sa neprekrývajúcich intervalov. Uvažované časové rady pozostávajú z 2641 položiek, ale na analýzu použijeme iba 2640, pretože takto dlhý rad môžeme rozdeliť do navzájom sa neprekrývajúcich období až 40 rôznych dĺžok. Pri R/S analýze sme použili iba dĺžky ≥ 10 , teda 33 rôznych dĺžok.

Hurstov koeficient sme odhadovali najskôr pre pôvodný rad výmenných kurzov. Výsledky sú v tabuľke 2. Vo všetkých štyroch prípadoch presiahli odhady hodnotu 0.95, čo poukazuje na veľmi silný trend. V prípade amerického dolára dokonca presiahol teoreticky najvyššiu možnú hodnotu 1. S takýmito výsledkami sa stretávame i v odbornej literatúre [14, 21]. Sú spôsobené nepresnosťami jednotlivých metód používaných na odhad Hurstovho koeficientu. Výsledky v prípade R/S analýzy sú do značnej miery ovplyvnené počtom údajov, z ktorých sa vychádza pri lineárnej regresii. Napriek určitej miere nepresnosti nám odhady Hurstovho koeficientu potvrdzujú prítomnosť trendu. A práve tento trend sa budeme v ďalšej časti snažiť modelovať.

	AUD	CAD	GBP	USD
Hurstov koeficient	0.95107	0.99279	0.98921	1.01974
štandardná odchýlka	0.015705	0.008144	0.004679	0.002562

Tabuľka 2: Odhady Hurstovho koeficientu pre výmenné kurzy

Výsledky získané odhadom Hurstovho koeficientu sú navyše v súlade s výsledkami získanými skúmaním autokorelačných funkcií. ACF v prípade amerického dolára odzrkadľuje najmenší pokles a teda najvyšší výskyt pamäťových znakov. Jeho Hurstov koeficient nadobúda najvyššiu hodnotu a teda poukazuje na prítomnosť výrazného trendu. Podobná úvaha platí aj v prípade austrálskeho dolára. Hurstov koeficient poukazuje na prítomnosť o niečo menej výrazného trendu. Tento stav odzrkadľuje autokorelogram výraznejším poklesom ACF.

	AUD	CAD	GBP	USD
Hurstov koeficient	0.50015	0.53880	0.55599	0.57633
štandardná odchýlka	0.011030	0.005612	0.006516	0.007500

Tabuľka 3: Odhady Hurstovho koeficientu pre výnosy výmenných kurzov

Ako sme už spomenuli, budeme modelovať nielen výmenné kurzy ale aj výnosy. Preto nás zaujímali aj ich hodnoty Hurstovho koeficientu. Keďže logaritmickou transformáciou cien sme vytvorili stacionárne časové rady, očakávali sme výrazne nižšie hodnoty pohybujúce sa nad úrovňou Hurstovho koeficientu pre náhodný proces, čo je 0.5. Výsledky sú zobrazené v tabuľke 3. Podobne aj v tomto prípade sme pre austrálsky dolár dostali najnižší odhad Hurstovho koeficientu, ktorý je však až nečakane blízko hodnote 0.5. Maximálny odhad Hurstovho koeficientu sme dosiahli rovnako ako v predchádzajúcom prípade pre americký dolár.

Na základe výsledkov R/S analýzy sme sa rozhodli predpovedať austrálsky dolár, ktorý má najnižšie hodnoty Hurstovho koeficientu, a americký dolár, ktorý nadobúda najvyššie hodnoty. Väčší koeficient amerického dolára indikuje lepší pamäťový efekt, ktorý by mal vyústiť do lepších výsledkov predpovedí. Naopak veľmi nízka hodnota v prípade výnosov austrálskeho dolára poukazuje na náhodný proces. Teda dnešná hodnota výnosu by nemala mať žiadny vplyv na jeho zajtrajšiu hodnotu.

Kapitola 6

Štúdia predikčnej schopnosti neurónových sietí

Na základe zistení uverejnených v piatej kapitole budeme predpovedať výmenné kurzy eura voči austrálskemu doláru a americkému doláru. Na predpovedanie použijeme dopredné neurónové siete tréňované metódou spätného šírenia chyby. I keď v priebehu tréňovania berieme za meradlo výkonnosti siete strednú kvadratickú chybu, výstupy budeme vyhodnocovať hlavne pomocou D_{stat} , S_{stat} a výnosov dosiahnutých trhovými simuláciami. Pri tréňovaní generujeme počiatkové váhy ako malé náhodné čísla, čo do značnej miery ovplyvňuje výsledok siete. Na dosiahnutie objektivity výsledkov budeme generovať 500 neurónových sietí, pričom za výsledok zoberieme priemernú hodnotu z jednotlivých výstupov. Na simuláciu neurónových sietí sme využívali nástroje programu MATLAB.

6.1. Návrh siete

Výsledky získané z neurónových sietí sú výrazne podmienené nielen voľbou učiaceho algoritmu, ale aj architektúrou siete a nastavením jednotlivých parametrov vstupujúcich do výpočtu. Ako učiaci algoritmus sme si v tejto práci zvolili metódu spätného šírenia chyby s náhodnou inicializáciou váh. Nastavenie ostatných parametrov siete popíšeme v nasledujúcej časti.

6.1.1. Príprava dát

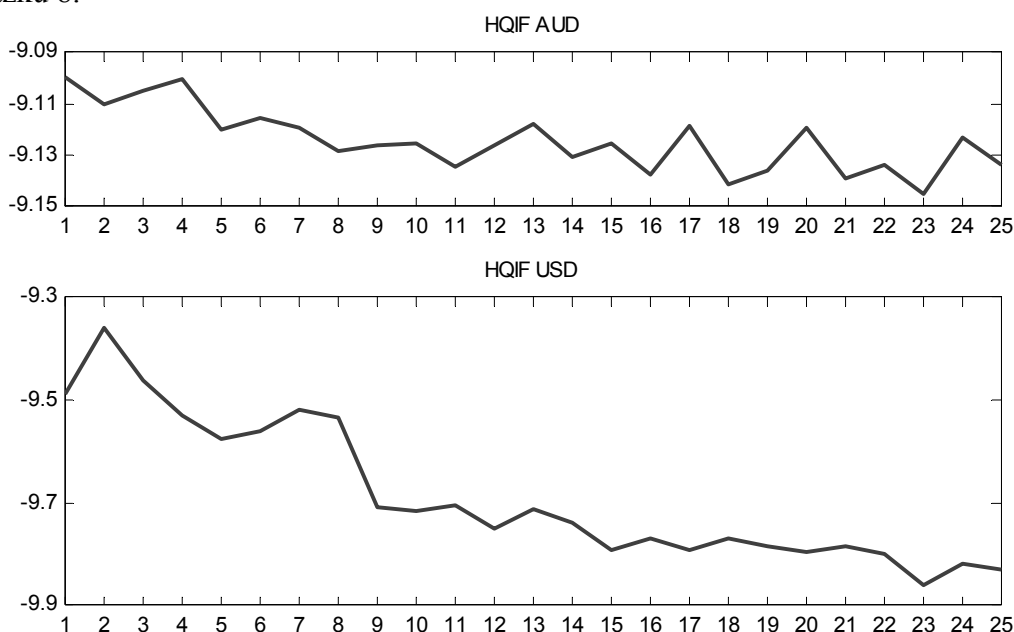
Analýzou výmenných kurzov sme zistili, že skúmané rady výmenných kurzov sú pomerne výrazne korelované. Budúce hodnoty by mali byť preto podmienené nielen súčasnými ale i minulými hodnotami. Aby sme boli schopní zachytiť tieto vzťahy pomocou dopredných neurónových sietí, vznikol prístup tzv. pohyblivého okna. Pôvodný rad je rozdelený do množín rovnakej dĺžky, ktoré tvoria vstup siete. Môžeme si to predstaviť ako posun šablóny po dátach, ktorá prechádza celou dátovou množinou a vytvára vstupno-výstupné páry. Vstupno-výstupné páry budú teda vyzeráť nasledovne

$$\begin{array}{ll} vstup = \{x_1, x_2, x_3, \dots, x_d\} & výstup = \{x_{d+1}\} \\ vstup = \{x_2, x_3, x_4, \dots, x_{d+1}\} & výstup = \{x_{d+2}\} \\ \vdots & \vdots \\ vstup = \{x_{2641-d}, x_{2641-(d+1)}, \dots, x_{2640}\} & výstup = \{x_{2641}\}. \end{array}$$

Ako vidíme, v tejto práci zameriame na predpovedanie hodnôt na nasledujúci pracovný deň, teda o 1 krok dopredu. Voľba rozmeru vstupných dát d je dôležitá pre úspešnosť predikcie. Autokorelačná funkcia výmenných kurzov odhalila silné korelácie i pri veľkých omeškaniach, avšak sieti by sme mali poskytnúť nejaké rozumné množstvo vstupov. Preto sme sa rozhodli použiť Hannan-Quinnovo informačné kritérium, ktoré penalizuje výkon siete počtom jej parametrov. Vypočítame ho podľa vzorca

$$HQIF = \ln \left[\frac{1}{N} \sum_{t=1}^N (y_t - \hat{y}_t)^2 \right] + \frac{k}{N} \ln[\ln(N)].$$

Budeme porovnávať hodnoty HQIF pre rozmery vstupných dát od 1 po 25. Vyberieme taký rozmer vstupov, ktorý bude nadobúdať čo najmenšie hodnoty HQIF pri rozumnej miere omeškania, aby sa minimalizoval počet vstupných neurónov a teda aj celkový čas tréningovania. Výsledky, ktoré sme dostali ako priemer z hodnôt pri 100 neurónových sieťach, sú zobrazené na obrázku 8.



Obrázok 8: Hodnoty Hannan-Quinn informačného kritéria pre rôzne rozmery vstupných vektorov

V prípade amerického dolára má funkcia klesajúci trend, zatiaľ čo pri austrálskom dolári sme svedkami pomerne výrazných výkyvov. Navyše rozdiel medzi maximálnou a minimálnou hodnotou Hannan-Quinn informačného kritéria predstavuje iba 0.0454, čo do istej miery podkopáva spoľahlivosť týchto výsledkov. Napriek tomu využijeme tieto informácie na určenie dimenzie vstupných dát. Pre AUD i USD nadobúda HQIF minimum v bode 23, čo sa nám však zdá ako príliš veľké dimenzia. Myslíme si, že vzhľadom na veľkosť siete by mala byť dimenzia vstupného vektora menšia ako 15. V prípade austrálskeho dolára nám potom vychádza 11 ako najvhodnejší rozmer vstupných dát. Pre americký dolár sme si zvolili dimenziu 12.

Pri výnosoch výmenných kurzov je situácia jednoduchšia. Rozmer vstupných dát určíme z grafov autokorelačnej funkcie. V prípade austrálskeho dolára nadobúda autokorelačný

koeficient výraznejšie hodnoty až po omeškanie 10, potom začnú mierne klesať. Vyššie hodnoty sa objavujú až pri omeškaniach 15 a 22. Preto v tomto prípade použijeme okno dĺžky 10. Pre americký dolár sú hodnoty autokorelačných koeficientov celkovo nižšie, vyššie hodnoty vidíme pri omeškaniach 9, 10 a 12. Preto v tomto prípade použijeme okno dĺžky 12.

Trénovacia, validačná a testovacia množina

Vytvorenú množinu vstupno-výstupných párov rozdelíme na tri časti – trénovaciú, validačnú a testovaciú množinu. Na trénovacej množine bude prebiehať učenie siete, teda cieľená postupná zmena synaptických váh. Validáciu množinu budeme používať na zabránenie pretrénovania siete. Testovaciú množinu použijeme vo fáze testovania siete, kedy na základe naučených váh získame príslušné predpovede. Dáta sme si rozdelili v nasledovnom pomere:

trénovacia množina	70%
validačná množina	20%
testovacia množina	10%

Tabuľka 4: Rozdelenie dát na trénovaciú, validačnú a testovaciú množinu

6.1.2. Výber architektúry

Počet nezávislých premenných, ktoré treba trénovať, ovplyvňuje dĺžku tréningu. Ak máme príliš veľkú architektúru a príliš veľa parametrov, výkon siete tým môžeme degradovať. Naopak, pri malom počte nedokážeme dostatočne opísať dáta. Kým počet vstupných a výstupných neurónov je možné určiť jednoznačne na základe dimenzie vstupných a výstupných vektorov, voľba počtu skrytých vrstiev, no najmä neurónov v rámci týchto vrstiev, zostala doposiaľ nevyriešeným problémom. Pri predpovedaní časových radov sa zvyčajne stretávame s jednou skrytou vrstvou, pričom počet neurónov sa líši v závislosti od konkrétneho problému. Bolo vypracovaných viacero štúdií, ktoré sa zaoberali optimálnou voľbou počtu skrytých vrstiev. V jednom z posledných publikovaných článkov [29] určili autori optimálny počet neurónov vzťahom

$$n = C(N/(d \log(N)))^{1/2},$$

kde C je konštanta, d je rozmer vstupných dát a N je počet trénovacích párov. Konštantu C určili tak, že pre postupnosť hodnôt C trénovali sieť s n skrytými neurónmi, pričom za optimálne vyberali to C , ktoré minimalizuje chybu siete. Na základe experimentov ukázali, že pre N/d väčšie ako 30 je optimálny počet skrytých neurónov približne $(N/(d \log(N)))^{1/2}$. Na určenie počtu neurónov v skrytej vrstve použijeme práve tento vzorec. Výsledné počty sú zobrazené v tabuľke 5.

	AUD	AUD výnosy	USD	USD výnosy
počet trénovacích párov	1839	1839	1839	1839
dimenzia vstupov	11	10	12	12
počet skrytých neurónov	5	5	5	5

Tabuľka 5: Počty skrytých neurónov pre jednotlivé vstupné premenné siete

6.1.3. Trénovací algoritmus

Pri tréovaní metódou spätného šírenia chyby môžeme na minimalizáciu chybovej funkcie použiť viacero metód. Budeme sa snažiť zistiť, do akej miery ovplyvňuje voľba metódy výsledok predikcie, a teda ktorá metóda je najvhodnejšia na predikciu. V teoretickej časti sme spomenuli tri metódy – gradientovú metódu, kvázinewtonovskú BFGS metódu a Levenberg-Marquardtovu metódu. My však kvôli vysokej výpočtovej náročnosti gradientovej metódy budeme používať iba zvyšné dve.

6.1.4. Zastavenie tréovania

Proces učenia bude prebiehať až do jeho zastavenia. Siete sa netrénujú až po dosiahnutie minimálnej chyby na tréovacej množine, keďže v tom prípade by prišlo k pretrénovaniu siete. Sieť by sa učila šum, čo je na úkor schopnosti zovšeobecňovania. Dôležitou úlohou je preto vystihnúť správny moment, v ktorom by malo byť tréovanie prerušené.

Neexistuje žiadne jednoznačné pravidlo na zastavenie učiaceho algoritmu. Existuje však niekoľko rozumných kritérií, ktoré môžu byť použité na ukončenie procesu učenia. V tejto práci budeme využívať tzv. koncept skorého zastavenia siete. Pri minimalizácii účelovej funkcie budeme sledovať aj chybu vo validačnej množine. V prípade, že chyba začne niekoľko cyklov po sebe rásť, proces tréovania sa ukončí. Za výsledné sa zoberú tie synaptické váhy, pri ktorých bola chyba vo validačnej množine minimálna. Dôležitá je pri tom nielen voľba validačnej množiny, ale aj maximálnej prípustnej miery zhoršenia na validačnej množine. Ich hodnoty sme určili experimentálne. Existuje viacero prístupov ku voľbe validačnej množiny, my sme však uvažovali nasledovné dva. Pod prístupom 1 budeme uvažovať usporiadanie dátovej množiny podľa času a jej následné rozdelenie na tréováciu a validačnú množinu v príslušnom pomere. Prístupom 2 budeme nazývať náhodné rozdelenie dát na validačné a testovacie avšak v príslušnom pomere. Uvažovali sme 5 rôznych validačných množín, ktoré kombinujú tieto dva prístupy. Ich popis je v tabuľke 6. Každú z týchto piatich možností sme testovali na 300 neurónových sieťach. Pri výbere výslednej validačnej množiny sme prihliadali najmä na priemerné hodnoty D_{stat} resp. S_{stat} , ale i na veľkosť MSE a výpočtovú náročnosť. Vybrané hodnoty sú zvýraznené hrubším písmom v tabuľke 7. Obdobným spôsobom, ako pri validačnej množine, sme postupovali aj pri maximálnej miere zhoršenia. V tomto prípade budeme vyberať najvhodnejšiu mieru zhoršenia zo 4 hodnôt, konkrétne pôjde o hodnoty 10, 20, 50 a 100. Výsledky sú uvedené v tabuľke 8.

	prístup 1	prístup 2
Validačná množina 1	0%	20%
Validačná množina 2	5%	15%
Validačná množina 3	10%	10%
Validačná množina 4	15%	5%
Validačná množina 5	20%	0%

Tabuľka 6: Zadefinovanie spôsobu vytvárania validačnej množiny

		USD			AUD		
		epochy	D _{stat}	MSE	epochy	D _{stat}	MSE
BFGS	Validačná množina 1	190.30	48.76	4.69E-05	140.78	50.97	1.12E-04
	Validačná množina 2	63.47	49.24	1.60E-04	149.39	50.88	1.11E-04
	Validačná množina 3	56.17	49.31	1.87E-04	125.88	50.91	1.13E-04
	Validačná množina 4	55.24	49.28	2.09E-04	139.85	50.94	1.12E-04
	Validačná množina 5	54.11	49.32	2.40E-04	133.89	50.94	1.13E-04
LM	Validačná množina 1	33.66	48.45	4.31E-05	31.02	51.86	1.05E-04
	Validačná množina 2	42.14	48.90	5.34E-05	31.03	51.19	1.07E-04
	Validačná množina 3	37.09	48.49	8.87E-05	31.77	51.81	1.05E-04
	Validačná množina 4	33.51	48.61	1.14E-04	32.59	51.74	1.05E-04
	Validačná množina 5	33.49	48.67	1.55E-04	33.75	51.81	1.05E-04

		USD výnosy			AUD výnosy		
		epochy	S _{stat}	MSE	epochy	S _{stat}	MSE
BFGS	Validačná množina 1	70.37	52.51	3.73E-05	71.75	53.44	3.94E-05
	Validačná množina 2	78.00	52.78	3.72E-05	73.87	53.67	3.92E-05
	Validačná množina 3	66.47	52.52	3.74E-05	74.79	53.67	3.93E-05
	Validačná množina 4	63.45	52.43	3.74E-05	72.27	53.62	3.93E-05
	Validačná množina 5	66.69	52.49	3.74E-05	79.35	53.94	3.91E-05
LM	Validačná množina 1	24.10	54.05	3.68E-05	24.05	54.06	3.90E-05
	Validačná množina 2	24.25	54.45	3.66E-05	23.95	53.97	3.94E-05
	Validačná množina 3	24.36	54.26	3.65E-05	23.86	53.85	3.92E-05
	Validačná množina 4	24.63	54.68	3.64E-05	24.09	53.84	3.90E-05
	Validačná množina 5	27.11	55.81	3.54E-05	24.24	54.43	3.87E-05

Tabuľka 7: Voľba validačnej množiny

		USD			AUD		
		epochy	D _{stat}	MSE	epochy	D _{stat}	MSE
BFGS	miera zhoršenia =10	29.69	49.48	7.54E-04	95.52	50.94	1.17E-04
	miera zhoršenia =20	56.58	49.29	5.05E-04	147.05	50.96	1.11E-04
	miera zhoršenia =50	121.85	49.02	1.19E-04	218.72	50.97	1.09E-04
	miera zhoršenia =100	216.30	48.81	7.55E-05	302.28	51.04	1.09E-04
LM	miera zhoršenia =10	21.46	48.62	1.34E-04	19.29	51.46	1.06E-04
	miera zhoršenia =20	33.47	48.68	1.45E-04	29.98	51.45	1.06E-04
	miera zhoršenia =50	68.47	48.73	1.21E-04	60.28	51.49	1.06E-04
	miera zhoršenia =100	127.95	48.87	1.25E-04	110.65	51.47	1.06E-04

		USD výnosy			AUD výnosy		
		epochy	S _{stat}	MSE	epochy	S _{stat}	MSE
BFGS	miera zhoršenia =10	45.61	52.28	3.77E-05	61.99	53.69	3.93E-05
	miera zhoršenia =20	65.91	52.49	3.74E-05	79.25	53.85	3.91E-05
	miera zhoršenia =50	117.83	52.85	3.71E-05	114.54	54.02	3.91E-05
	miera zhoršenia =100	197.66	53.34	3.67E-05	165.77	54.14	3.90E-05
LM	miera zhoršenia =10	16.62	55.77	3.55E-05	14.15	54.25	3.88E-05
	miera zhoršenia =20	27.23	55.85	3.53E-05	24.20	54.34	3.87E-05
	miera zhoršenia =50	56.86	55.74	3.54E-05	54.29	54.33	3.87E-05
	miera zhoršenia =100	107.20	55.85	3.54E-05	104.31	54.34	3.88E-05

Tabuľka 8: Voľba maximálnej miery zhoršenia na validačnej množine

6.2. Analýza predikčnej schopnosti neurónových sietí

Po tom, čo sme zvolili nastavenie siete, sme pristúpili ku simulácií a následnému predpovedaniu výmenných kurzov. Začali sme predpovedaním samotných hodnôt výmenných kurzov, neskôr sme prešli ku predpovedaniu výnosov.

6.2.1. Predpovedanie výmenných kurzov

Nasledujúca tabuľka obsahuje priemerné hodnoty výstupov, ktoré sme získali simuláciou 500 neurónových sietí.

		AUD		USD		
		BFGS	LM	BFGS	LM	
	epochy	258.97	67.14	33.39	166.06	
Simulácia	D _{stat}	51.14	51.61	49.54	49.35	
	MSE	1.11E-04	1.07E-04	5.75E-04	7.47E-05	
	D _{stat}	44.98	45.07	52.05	53.67	
Predpoveď	MSE	2.20E-03	4.03E-03	5.10E-03	5.21E-03	
	počet transakcií na trhu	148.35	151.03	129.75	134.39	
	bez nákladov	zisk	-48.13	-64.23	121.77	180.84
Stratégia 1	výnos	-4.76	-6.35	12.03	17.86	
	s nákladmi	zisk	-343.19	-358.55	-187.77	-155.30
	výnos	-34.00	-35.53	-18.60	-15.40	
Stratégia 2	počet transakcií na trhu	87.68	92.30	38.97	54.30	
	bez nákladov	zisk	-118.00	-139.59	81.76	226.20
	výnos	-11.68	-13.81	8.08	22.33	
Stratégia 2	s nákladmi	zisk	-291.20	-316.14	-17.26	72.85
	výnos	-28.84	-31.32	-1.71	7.21	

Tabuľka 9: Priemerné výsledky predpovedí

V prípade austrálskeho dolára vidíme, že simuláciou na tréningových a validačných dátach dosahujeme úspešnosť len niečo cez 51 %. Úspešnosť predpovede je ešte nižšia, pohybuje sa na úrovni 45 %. To platí ako pre Levenberg-Marquardtovu tak aj pre BFGS metódu, ktoré dávajú približne rovnaké hodnoty D_{stat} aj MSE. Trhovou simuláciou obchodných stratégií tiež dosiahneme porovnateľné výsledky. V prípade BFGS metódy je zisk pri oboch stratégiách o niečo vyšší. Platí to pre stratégií s nákladmi aj bez zahrnutia nákladov.

V prípade amerického dolára získame simuláciou dát síce nižšiu úspešnosť, iba na úrovni 49 %. Pri predpovedaní však stúpne až na úroveň 52-53 %. Oproti austrálskemu doláru vidíme výrazné rozdiely medzi jednotlivými metódami. Levenberg-Marquardtovou metódou dosahujeme pri oboch stratégiách výrazne vyšší zisk ako pri BFGS metóde. Pri druhej stratégií dosahujeme zisk aj po zahrnutí nákladov.

Okrem priemerných výsledkov nás zaujímali aj maximálne hodnoty, resp. maximálny výnos, ktorý by sme mohli dosiahnuť. Maximálne hodnoty pre každú jednu položku z tabuľky 9 sú zobrazené v tabuľke 10. Vidíme, že aj v prípade austrálskeho dolára by sme boli schopní dosiahnuť zisk, i keď len bez zahrnutia transakčných nákladov. Maximálne hodnoty sú do značnej miery podmienené voľbou metódy, čo sa však v priemere neukázalo. Pri americkom doláre sa maximálna hodnota úspešnosti predpovede blíži ku 60 %, čo sa

prejavilo na vysokom zisku. V kladných číslach sa pri tom pohybujeme bez ohľadu na použitú minimalizačnú metódu a obchodnú stratégiu. Pri stratégii 1 bol maximálny zisk vyčíslený použitím BFGS metódy, ktorá však v priemere skončila horšie ako Levenberg-Marquardtova metóda.

			AUD		USD	
			BFGS	LM	BFGS	LM
Stratégia 1	Simulácia	D _{stat}	52.22	53.02	52.12	52.58
	Predpoveď	D _{stat}	52.47	49.81	59.32	59.70
	bez nákladov	zisk	114.49	141.96	390.59	345.40
		výnos	11.31	14.02	38.54	34.09
	s nákladmi	zisk	-214.98	-223.33	112.60	8.85
		výnos	-21.28	-22.11	11.15	0.88
Stratégia 2	bez nákladov	zisk	97.77	48.67	341.51	429.47
		výnos	9.66	4.81	33.70	42.37
	s nákladmi	zisk	-130.21	-148.29	276.90	320.46
		výnos	-12.88	-14.67	27.33	31.71

Tabuľka 10: Najlepšie výsledky predpovedí

Tieto vysoké hodnoty však nepochádzajú z jednej simulácie. To nás privádza ku otázke, na základe akého kritéria vybrať sieť, ktorá nám dá čo možno najziskovejšiu predpoveď. Ku možným kritériám patrí najvyššia hodnota D_{stat} alebo najnižšia hodnota MSE. To, že tieto dva kritériá nie sú zárukou vysokého zisku, vidíme z tabuľky 11. Pre každý výmenný kurz sú tam zobrazené výsledky štyroch konkrétnych simulácií pomocou Levenberg-Marquardtovej metódy. Je tam simulácia s najvyššou D_{stat}, najnižšou MSE, najvyšším ziskom pri stratégii 1 a najvyšším ziskom pri stratégii 2. Pre USD je najvyššia dosiahnutá úspešnosť 52.58%. Tá sa síce odrazila v nadpriemernom zisku, no zďaleka nie najvyššom. Najvyšší zisk bol dosiahnutý v prípadoch, kedy bola úspešnosť simulácie rádovo o 2% nižšia. Pri simulácií obchodnej stratégie 2 sieťou s najnižšou MSE sme dostali dokonca podpriemerné výsledky. Pri AUD vidíme, že najvyššia úspešnosť a najnižšia MSE boli dosiahnuté pri tej istej simulácii. Ani táto kombinácia nezaručila dobré výsledky. Dokonca sú podpriemerné. Maximálny zisk bol znova dosiahnutý simuláciami, ktoré mali úspešnosť o 1-2% nižšiu.

			AUD				USD			
Simulácia	D _{stat}	53.02	53.02	51.80	51.21	52.58	50.76	49.15	49.28	
	MSE	1.01E-04	1.01E-04	1.06E-04	1.08E-04	7.81E-05	4.15E-05	5.27E-05	8.06E-05	
Predpoveď	D _{stat}	45.25	45.25	49.05	44.49	55.51	55.13	57.41	55.51	
	MSE	1.72E-02	1.72E-02	3.56E-03	1.70E-03	1.75E-03	3.07E-04	8.63E-04	2.88E-03	
S1	bez nákladov	zisk	-111.89	-111.89	141.96	-113.70	223.47	194.54	345.40	222.63
	s nákladmi	zisk	-398.99	-398.99	-223.33	-400.22	-103.00	-150.12	8.85	-127.64
S2	bez nákladov	zisk	-222.30	-222.30	-106.82	48.67	276.11	185.59	240.96	429.47
	s nákladmi	zisk	-435.50	-435.50	-299.37	-227.30	145.89	-48.81	130.86	238.80

Tabuľka 11: Štyri konkrétne predpovede

6.2.2. Predpovedanie výnosov výmenných kurzov

Rovnaký postup ako pri výmenných kurzoch sme aplikovali aj na ich výnosy. Výsledky sú zobrazené v tabuľkách 12, 13 a 14.

			AUD		USD	
			BFGS	LM	BFGS	LM
Stratégia 1		epochy	159.55	104.86	171.56	109.40
	Simulácia	S _{stat}	53.95	53.79	52.81	54.21
		MSE	3.90E-05	3.89E-05	3.69E-05	3.62E-05
	Predpoveď	S _{stat}	48.78	48.67	50.19	49.86
		MSE	2.02E-03	1.71E-04	5.43E-04	1.71E-04
		počet transakcií na trhu	147.58	129.46	99.69	99.34
	bez nákladov	zisk	-59.22	-76.98	101.86	100.13
		výnos	-5.84	-7.59	10.02	9.85
	s nákladmi	zisk	-348.39	-329.92	-140.51	-141.65
		výnos	-34.50	-32.67	-13.88	-13.99

Tabuľka 12: Priemerné výsledky predpovedí

Pre oba rady výnosov sme simuláciou dosiahli vyššiu úspešnosť, ako tomu bolo v prípade samotných výmenných kurzov. Pre austrálsky dolár je úspešnosť predpovede rádovo o 3 % vyššia. Opačná situácia nastala v prípade amerického dolára, kde je úspešnosť predpovede o 2-3 % nižšia. V oboch prípadoch je však zisk získaný simuláciou obchodnej stratégie nižší ako pri samotnom výmennom kurze.

V tabuľke 13 sú zachytené maximálne hodnoty jednotlivých veličín získaných z predpovedí. Znovu vidíme pomerne vysoké čísla, ktoré pre americký dolár aj pri zahrnutí nákladov vykazujú zisk.

			AUD		USD	
			BFGS	LM	BFGS	LM
Stratégia 1	Simulácia	S _{stat}	56.93	58.79	55.67	58.04
	Predpoveď	S _{stat}	57.58	56.06	58.71	57.58
	bez nákladov	zisk	126.16	210.97	367.77	364.56
		výnos	12.41	20.75	36.13	25.36
	s nákladmi	zisk	-192.47	-28.44	49.77	35.81
		výnos	-19.03	-2.80	4.90	2.50

Tabuľka 13: Najlepšie výsledky predpovedí

Aj v prípade výnosov sme svedkom toho, že vysoká úspešnosť simulácie alebo nízka hodnota MSE nie sú zárukou vysokého zisku. Najvyšší zisk sme dosiahli v prípadoch, kedy bola úspešnosť simulácie podpriemerná.

			AUD			USD		
Simulácia	S _{stat}	58.79	57.52	52.75	58.04	55.46	53.55	
	MSE	3.71E-05	3.59E-05	3.91E-05	3.53E-05	3.33E-05	3.60E-05	
Predpoveď	S _{stat}	47.35	49.62	56.06	55.68	47.35	53.79	
	MSE	1.72E-04	1.84E-04	1.77E-04	1.34E-04	3.83E-04	9.41E-05	
S1	bez nákladov	zisk	-75.93	-132.06	210.97	235.83	134.92	364.56
	s nákladmi	zisk	-333.44	-415.99	-28.44	-91.43	-80.29	-9.50

Tabuľka 14: Tri konkrétne predpovede

6.3. Porovnanie klasických neurónových sietí a neurónových sietí druhého rádu

Na predpovedanie výmenných kurzov budeme používať nielen klasické neurónové siete, ale aj neurónové siete vyššieho rádu. Aby sme ich vedeli porovnať, aplikovali sme nasledovné dva prístupy. V prvom prípade sme použili 4 vstupné údaje pri klasickej neurónovej sieti i neurónovej sieti druhého rádu. V prípade HONN to však predstavovalo po zahrnutí násobných členov spolu až 10 vstupných neurónov. V druhom prípade sme klasickej neurónovej sieti zadali vstup dĺžky 10, pričom HONN dostala znovu iba 4 vstupy.

6.3.1. Porovnanie pre výmenné kurzy

		BFGS metóda		LM metóda			
		ANN	HONN	ANN	HONN		
Stratégia 1	Simulácia	epochy	30.67	33.71	157.52	123.07	
		D ^{stat}	49.48	49.27	48.78	48.70	
		MSE	9.18E-04	4.79E-04	1.00E-04	1.43E-04	
	Predpoveď	D ^{stat}	51.86	52.75	54.02	53.70	
		MSE	7.70E-03	4.48E-03	2.62E-03	2.86E-03	
	bez nákladov	zisk	127.95	148.36	175.47	166.77	
		výnos	12.59	14.59	17.26	16.40	
	s nákladmi	zisk	-195.15	-147.66	-158.17	-162.78	
		výnos	-19.28	-14.58	-15.64	-16.09	
	Stratégia 2	bez nákladov	zisk	97.60	79.40	200.72	200.84
			výnos	9.60	7.81	19.74	19.75
		s nákladmi	zisk	-9.98	-39.23	71.65	94.72
			výnos	-0.99	-3.87	7.06	9.34

Tabuľka 15: Porovnanie ANN a HONN pre výmenné kurzy USD, ANN so 4 vstupmi, HONN druhého rádu so 4 vstupmi, tj. spolu 10 vstupov

		BFGS metóda		LM metóda			
		ANN	HONN	ANN	HONN		
Stratégia 1	Simulácia	epochy	32.97	33.68	161.18	125.62	
		D ^{stat}	49.46	49.18	48.64	48.69	
		MSE	5.29E-04	5.49E-04	6.69E-05	1.28E-04	
	Predpoveď	D ^{stat}	51.62	52.71	53.77	53.57	
		MSE	4.93E-03	5.10E-03	1.54E-03	2.52E-03	
	bez nákladov	zisk	111.36	144.63	173.06	164.44	
		výnos	10.95	14.23	17.02	16.18	
	s nákladmi	zisk	-202.28	-150.50	-157.29	-165.82	
		výnos	-19.98	-14.86	-15.55	-16.39	
	Stratégia 2	bez nákladov	zisk	78.39	85.61	222.05	204.65
			výnos	7.71	8.42	21.83	20.12
		s nákladmi	zisk	-24.72	-30.60	90.16	94.66
			výnos	-2.44	-3.02	8.89	9.34

Tabuľka 16: Porovnanie ANN a HONN pre výmenné kurzy USD, ANN s 10 vstupmi, HONN druhého rádu so 4 vstupmi, tj. spolu 10 vstupov

Tabuľky 15 a 16 obsahujú vyhodnotenie predpovedí kurzu eura voči americkému doláru. Vidíme, že výsledky získané pomocou ANN a HONN sa líšia v závislosti od meny a použitej metódy. V istých prípadoch sa ukazujú lepšie ANN, v iných zase HONN. I keď sú výsledky

v niektorých prípadoch pomerne tesné, vidíme v tabuľkách nasledovnú pravidelnosť: Pri použití BFGS metódy sú na simuláciu stratégie 1 výhodnejšie HONN, na simuláciu stratégie 2 zase ANN. Pri Levenberg-Marquardtovej metóde vidíme presne opačný výsledok, teda na simuláciu stratégie 1 sú výhodnejšie ANN, na stratégiu 2 zase HONN.

			BFGS metóda		LM metóda	
			ANN	HONN	ANN	HONN
Stratégia 2 Stratégia 1	Simulácia	epochy	216.72	194.88	66.88	64.06
		D _{stat}	50.61	50.61	50.86	50.85
		MSE	1.11E-04	1.11E-04	1.11E-04	1.11E-04
	Predpoveď	D _{stat}	45.13	45.08	45.55	45.83
		MSE	2.16E-03	1.90E-03	3.61E-03	1.85E-02
	bez nákladov	zisk	-59.12	-48.30	-59.18	-76.02
		výnos	-5.83	-4.76	-5.83	-7.49
	s nákladmi	zisk	-350.08	-341.28	-354.47	-365.09
		výnos	-34.58	-33.71	-35.02	-36.07
	bez nákladov	zisk	-121.41	-121.58	-110.98	-89.18
		výnos	-11.97	-11.99	-10.94	-8.79
	s nákladmi	zisk	-283.41	-287.50	-307.83	-279.46
		výnos	-27.98	-28.38	-30.40	-27.59

Tabuľka 17: Porovnanie ANN a HONN pre výmenné kurzy AUD, ANN so 4 vstupmi, HONN druhého rádu so 4 vstupmi, tj. spolu 10 vstupov

			BFGS metóda		LM metóda	
			ANN	HONN	ANN	HONN
Stratégia 2 Stratégia 1	Simulácia	epochy	278.55	190.15	62.25	70.77
		D _{stat}	50.93	50.57	51.19	50.77
		MSE	1.09E-04	1.11E-04	1.08E-04	1.10E-04
	Predpoveď	D _{stat}	45.32	45.23	46.50	44.98
		MSE	2.34E-03	1.72E-03	2.93E-03	3.82E-03
	bez nákladov	zisk	-44.98	-44.45	-34.41	-76.35
		výnos	-4.43	-4.38	-3.39	-7.53
	s nákladmi	zisk	-341.25	-337.41	-337.94	-360.05
		výnos	-33.71	-33.33	-33.38	-35.57
	bez nákladov	zisk	-110.76	-124.65	-113.37	-116.80
		výnos	-10.92	-12.29	-11.18	-11.51
	s nákladmi	zisk	-275.01	-292.13	-285.29	-277.62
		výnos	-27.15	-28.84	-28.17	-27.41

Tabuľka 18: Porovnanie ANN a HONN pre výmenné kurzy AUD, ANN s 10 vstupmi, HONN druhého rádu so 4 vstupmi, tj. spolu 10 vstupov

Výsledky dosiahnuté predpovedaním výmenného kurzu austrálskeho dolára sú do istej miery podobné s výsledkami získanými pri americkom dolári. Výsledky majú presne rovnaký zrkadlový charakter. Teda pri použití BFGS metódy sú na simuláciu stratégie 1 výhodnejšie HONN, na stratégiu 2 zase ANN. Pri Levenberg-Marquardtovej metóde sú naopak na simuláciu stratégie 1 výhodnejšie ANN, na stratégiu 2 zase HONN.

6.3.2. Porovnanie pre výnosy výmenných kurzov

Rovnaký postup sme aplikovali aj na výnosy výmenných kurzov. Výsledky pre výmenný kurz eura voči americkému doláru sú zobrazené v tabuľkách 19 a 20. Vidíme, že ANN i HONN takmer rovnaké hodnoty S_{stat} a MSE. Simuláciou obchodnej stratégie 1 dosahujeme

podobné výsledky ako pri predpovedaní samotných výmenných kurzov, teda pri Levenberg-Marquardtovej metóde je lepšie ANN, pri BFGS metóde je lepšie HONN.

			BFG metóda		LM metóda	
			ANN	HONN	ANN	HONN
Stratégia 1	Simulácia	epochy	155.86	145.24	113.27	114.01
		S _{stat}	52.37	52.55	52.99	53.49
		MSE	3.73E-05	3.72E-05	3.69E-05	3.63E-05
	Predpoveď	S _{stat}	48.65	48.49	48.18	45.83
		MSE	3.99E-01	1.52E-04	1.17E-04	1.43E-04
	bez nákladov	zisk	59.48	71.27	65.08	61.78
		výnos	5.85	7.01	6.40	6.08
	s nákladmi	zisk	-165.69	-160.20	-158.13	-171.80
		výnos	-16.35	-15.80	-15.60	-16.95

Tabuľka 19: Porovnanie ANN a HONN pre výnosy výmenných kurzov USD, ANN so 4 vstupmi, HONN druhého rádu so 4 vstupmi, tj. spolu 10 vstupov

			BFGS metóda		LM metóda	
			ANN	HONN	ANN	HONN
Stratégia 1	Simulácia	epochy	172.65	142.28	109.21	112.21
		S _{stat}	52.58	52.46	53.71	53.56
		MSE	3.70E-05	3.72E-05	3.64E-05	3.63E-05
	Predpoveď	S _{stat}	48.77	48.24	48.70	48.06
		MSE	5.67E-01	1.03E-04	1.60E-04	1.39E-04
	bez nákladov	zisk	62.88	68.90	76.99	58.92
		výnos	6.19	6.78	7.57	5.80
	s nákladmi	zisk	-151.78	-163.10	-149.05	-170.95
		výnos	-14.99	-16.09	-14.72	-16.87

Tabuľka 20: Porovnanie ANN a HONN pre výnosy výmenných kurzov USD, ANN s 10 vstupmi, HONN druhého rádu so 4 vstupmi, tj. spolu 10 vstupov

Z tabuľky 21 vidíme, že pri simulácií vykazujú ANN a HONN takmer rovnaké hodnoty S_{stat} a MSE. Úspešnosť predpovede je však vyššia pri HONN, čo sa premietlo do vyššieho zisku. A to bez ohľadu na použitú minimalizačnú metódu. Z výsledkov zobrazených v tabuľke 22 je zrejmé, že vyššiu úspešnosť simulácie sme dosiahli pomocou ANN. Napriek tomu sú výnosy vyššie opäť pri neuronových sieťach druhého rádu.

			BFGS metóda		LM metóda	
			ANN	HONN	ANN	HONN
Stratégia 1	Simulácia	epochy	148.51	157.00	105.50	105.37
		S _{stat}	53.16	53.10	52.97	52.98
		MSE	3.98E-05	3.97E-05	3.98E-05	3.97E-05
	Predpoveď	S _{stat}	46.44	48.20	47.04	48.57
		MSE	2.01E-04	1.65E-03	1.68E-04	1.91E-04
	bez nákladov	zisk	-69.80	-28.67	-75.50	-35.13
		výnos	-6.88	-2.83	-7.44	-3.47
	s nákladmi	zisk	-328.25	-296.04	-307.34	-259.76
		výnos	-32.50	-29.30	-30.42	-25.70

Tabuľka 21: Porovnanie ANN a HONN pre výnosy výmenných kurzov AUD, ANN so 4 vstupmi, HONN druhého rádu so 4 vstupmi, tj. spolu 10 vstupov

			BFGS metóda		LM metóda	
			ANN	HONN	ANN	HONN
Stratégia 1	Simulácia	epochy	158.16	161.06	104.69	105.46
		S _{stat}	54.06	53.09	53.90	53.06
		MSE	3.90E-05	3.96E-05	3.89E-05	3.97E-05
	Predpoveď	S _{stat}	48.61	48.05	48.46	48.53
		MSE	5.91E-04	4.84E-04	1.71E-04	1.88E-04
	bez nákladov	zisk	-57.52	-35.48	-73.94	-37.26
		výnos	-5.67	-3.50	-7.29	-3.68
	s nákladmi	zisk	-343.88	-301.78	-331.83	-261.79
		výnos	-34.05	-29.87	-32.85	-25.90

Tabuľka 22: Porovnanie ANN a HONN pre výnosy výmenných kurzov AUD, ANN s 10 vstupmi, HONN druhého rádu so 4 vstupmi, tj. spolu 10 vstupov

6.4. Porovnanie predikčnej schopnosti neurónových sietí s inými metódami

Ako sme už spomenuli, kľúčovým momentom predpovedania je vyhodnotenie predikčnej schopnosti metódy. Preto sme sa rozhodli porovnať výsledky získané neurónovými sieťami s výsledkami, ktoré sme schopní dosiahnuť inými metódami predpovedania. V prvom rade pôjde o porovnanie s klasickými štatistickými metódami, teda modelom ARMA - GARCH. Potom pristúpime ku porovnaniu s naivnou predikciou.

6.4.1. Porovnanie predpovede s ARMA - GARCH modelom

Na základe štatistickej analýzy dát sme dospeli k záveru, že aplikovať klasický ARMA proces na predpovedanie výnosov by nevedlo ku uspokojuvým výsledkom. Museli by sme predpokladať konštantnú volatilitu, čo pri finančných časových radoch nemá opodstatnenie. Preto sme sa rozhodli modelovať aj volatilitu výnosov. Pre obe meny budeme hľadať ten najvhodnejší model, pričom budeme vychádzať z nasledovných typov modelov: ARCH, GARCH, ARMA + ARCH a ARMA + GARCH. Výsledky získané aplikovaním modelu ako klasickej štatistickej metódy porovnáme s výsledkami získanými pomocou neurónových sietí. Pri teoretickom pojednaní o GARCH modeloch sme vychádzali z toho, že podmienené rozdelenie chýb je normálne. V tomto tvare bol formulovaný aj pôvodný model. Odvtedy už bolo vypracovaných viacero štúdií, ktoré modifikujú pôvodný model použitím alternatívnych rozdelení. Najčastejšie ide o Studentovo t-rozdelenie alebo GED rozdelenie (z angl. *Generalized Error Distribution*). Výber rozdelenia je pritom podmienený konkrétnymi dátami. Existuje niekoľko metód, ktorými by sme mohli určiť typ rozdelenia, z ktorého pochádzajú dáta. Jednou z nich je Q-Q plot. Ten v oboch prípadoch naznačuje, že rezíduá nie sú normálne rozdelené, čo potvrdili aj štatistické testy. Pri americkom doláre sme sa rozhodli vychádzať z GED rozdelenia. Výnosy austrálskeho dolára dosahujú pomerne vysoké hodnoty špicatosti až 9.272207. Keďže Studentovo t-rozdelenie dokáže lepšie zachytiť špicatosť výnosov [28], použijeme namiesto normálneho rozdelenia práve Studentovo t-rozdelenie.

Pri modelovaní amerického dolára sme začali ARCH(1) modelom. Ljung-Boxovým testom sme overovali nulovú hypotézu o nezávislosti radov štandardizovaných rezíduí a ich druhých mocnín. V prípade rezíduí test síce nezamietol nulovú hypotézu, no ich druhé mocniny

označil ako nenáhodné. Preto sme na tento rad aplikovali GARCH(1,1) model. Vtedy už test nezamietol nulovú hypotézu aj pre druhé mocniny štandardizovaných rezíduí. Výsledky modelovania výnosov amerického dolára GARCH(1,1) modelom sú uvedené v nasledujúcej tabuľke.

	Odhad	Št. odchýlka	t-hodnota	Pr(> t)
mu	1.55E-04	1.12E-04	1.389	0.165
omega	8.89E-08	5.55E-08	1.602	0.109
alpha1	0.021	4.08E-03	5.243	1.58E-07 ***
beta1	0.976	4.31E-03	226.5	< 2e-16 ***
shape	1.542	0.067	23.117	< 2e-16 ***

Testy rezíduí			Štatistika	p-hodnota
Jarque-Bera Test	R	Chi ²	108.503	0
Shapiro-Wilk Test	R	W	0.993	2.44E-09
Ljung-Box Test	R	Q(10)	3.061	0.980
Ljung-Box Test	R	Q(15)	8.770	0.889
Ljung-Box Test	R	Q(20)	12.177	0.910
Ljung-Box Test	R ²	Q(10)	8.100	0.619
Ljung-Box Test	R ²	Q(15)	10.854	0.763
Ljung-Box Test	R ²	Q(20)	24.575	0.218

Tabuľka 23: Výsledky GARCH(1,1) modelovania výnosov USD

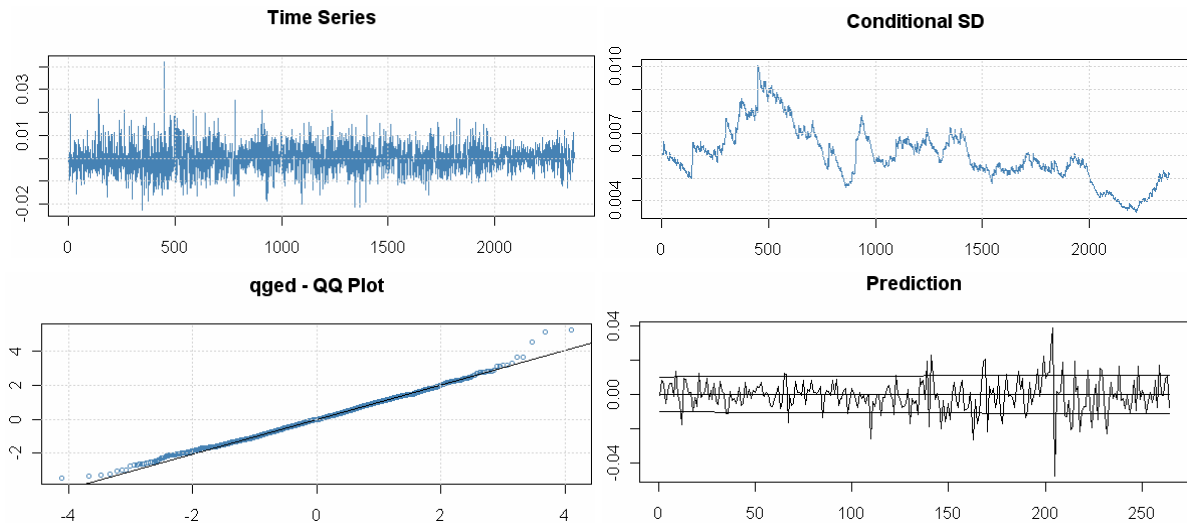
V prípade austrálskeho dolára je situácia trochu komplikovanejšia. Pri ARCH i GARCH modeli Ljung-Boxov test pomerne výrazne zamietol náhodnosť rezíduí. Preto sme pristúpili ku vytvoreniu kombinovaného ARMA + GARCH modelu. Výsledný model, ktorým sme sa rozhodli modelovať výnosy, má tvar ARMA(1,2) + GARCH (1,1). Výsledky sú zobrazené v tabuľke 24. Vidíme, že Ljung-Boxov test nezamieta nulovú hypotézu pri rezíduách len pre počty omeškania 10 a 20. Navyše nie veľmi výrazne. Pri druhých mocninách rezíduí nezamieta náhodnosť len pre 10 omeškaní.

	Odhad	Št. odchýlka	t-hodnota	Pr(> t)
mu	-7.57E-06	4.90E-06	-1.546	0.122
ar1	0.955	1.66E-02	57.616	< 2e-16 ***
ma1	-0.900	0.026	-34.499	< 2e-16 ***
ma2	-0.069	0.021	-3.367	7.60E-04 ***
omega	1.92E-07	9.05E-08	2.119	0.034 *
alpha1	0.039	7.46E-03	5.197	2.03E-07 ***
beta1	0.957	7.99E-03	119.770	< 2e-16 ***
shape	5.998	0.718	8.354	< 2e-16 ***

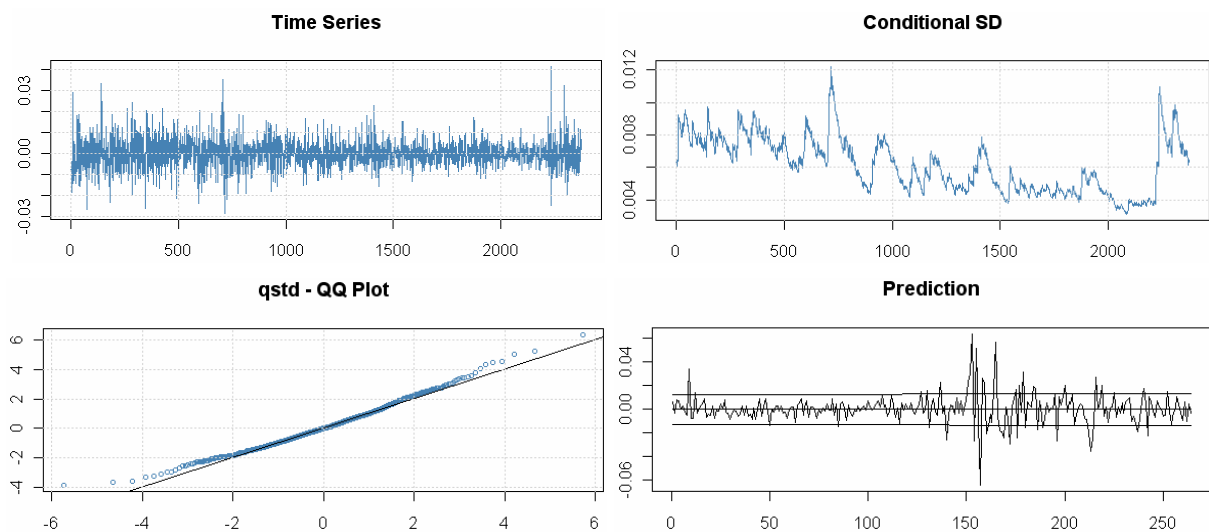
Testy rezíduí			Štatistika	p-hodnota
Jarque-Bera Test	R	Chi ²	579.733	0
Shapiro-Wilk Test	R	W	0.978	0
Ljung-Box Test	R	Q(10)	13.951	0.175
Ljung-Box Test	R	Q(15)	26.210	0.036
Ljung-Box Test	R	Q(20)	30.911	0.056
Ljung-Box Test	R ²	Q(10)	16.682	0.082
Ljung-Box Test	R ²	Q(15)	39.882	4.72E-04
Ljung-Box Test	R ²	Q(20)	43.044	2.02E-03

Tabuľka 24: Výsledky ARMA(1,2)+GARCH(1,1) modelovania výnosov AUD

Ak budeme takto vytvorené modely aplikovať na simulovanie a predpovedanie volatility, dostaneme hodnoty znázornené na obrázkoch 9 a 10. Priemerná predpovedaná hodnota volatility sa pre austrálsky dolár pohybuje na úrovni 0.006503707, pre americký dolár je to o niečo menej, konkrétne 0.006295158. Obe pritom s časom rastú.

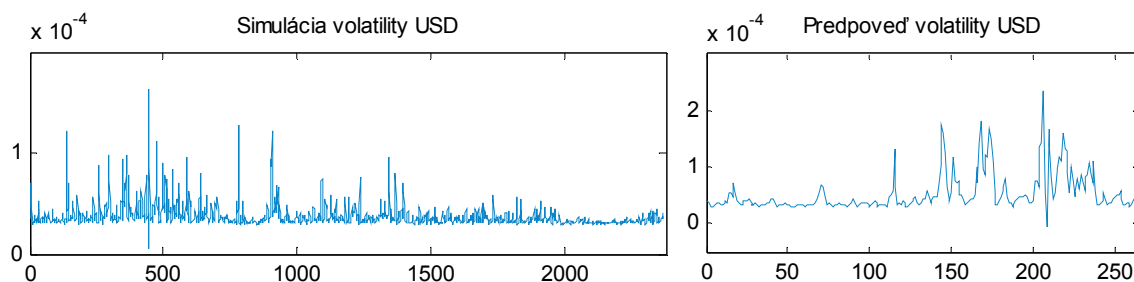


Obrázok 9: Grafické znázornenie volatility pre americký dolár

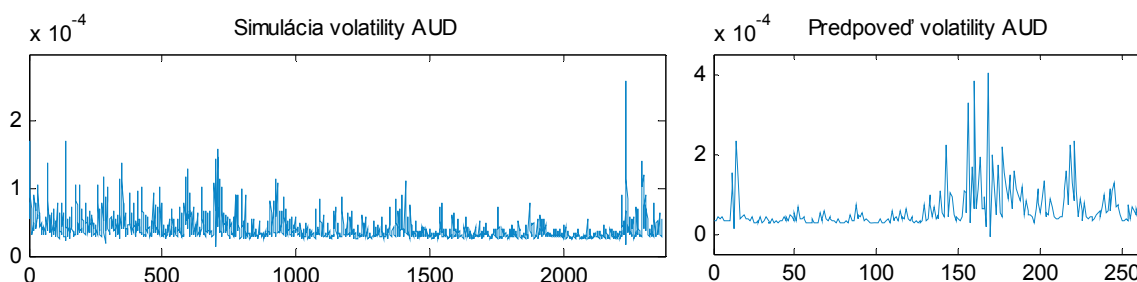


Obrázok 10: Grafické znázornenie volatility pre austrálsky dolár

Teraz sa pozrime na výsledky, ktoré sme schopní dosiahnuť pomocou neurónových sietí. Spôsob výpočtu volatility pomocou neurónových sietí je popísaný v prílohe C. Výsledky, ktoré sme dostali, sú zobrazené na obrázkoch 11 a 12. Vidíme, že predpovedaná hodnota má kvalitatívne úplne iný charakter. Pri klasickom prístupe na základe modelu odhadneme volatility na 264 dní dopredu. Pri neurónových sieťach ide o 264 predpovedí na o jeden časový krok dopredu. Keďže ide o dva rôzne prístupy, ťažko sa porovnávajú. Priemerná predpovedaná hodnota volatility sa pre austrálsky dolár sa pohybuje na úrovni 0.007818631719789, pre americký dolár je to o niečo menej, konkrétne 0.007294173611468. Tieto výsledky sú na porovnateľnej úrovni s tými, ktoré sme dostali klasickým prístupom.



Obrázok 11: Grafické znázornenie volatility pre americký dolár



Obrázok 12: Grafické znázornenie volatility pre austrálsky dolár

6.4.2. Porovnanie kvality predpovede s alternatívnymi metódami

Ako sme už spomenuli, kľúčovým momentom predpovedania je vyhodnotenie predikčnej schopnosti metódy. Výsledky prezentované v predchádzajúcich častiach porovnáme s výsledkami, ktoré sme schopní dosiahnuť alternatívnymi metódami. V prvom rade pôjde o výnosy dosiahnuté naivnou predikciou. Aj v tomto prípade budeme uvažovať obchodnú stratégiu s nákladmi i bez nákladov. Ďalej nás budú zaujímať výšky výnosov, ktoré by sme dosiahli pri klasickej Buy & Hold stratégii. Ročné výnosy spolu s výškami ziskov sú zobrazené v tabuľke 25. Ich priebeh za celé uvažované obdobie je graficky zobrazený na obrázku 13.

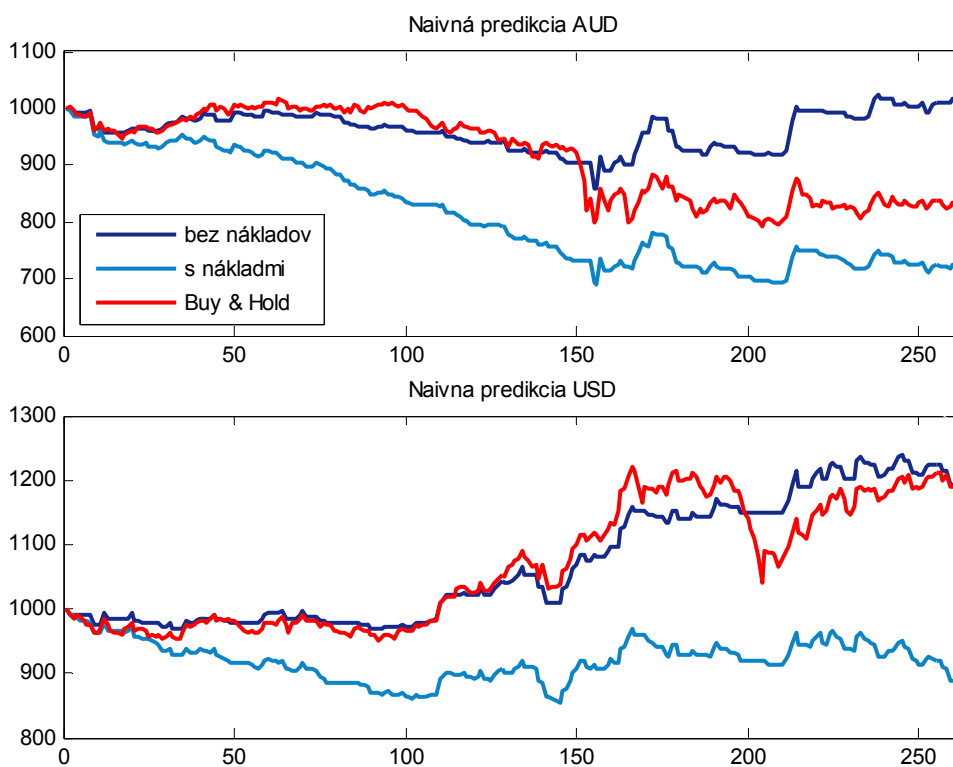
	Naivná predikcia				Buy & Hold	
	bez nákladov		s nákladmi		zisk	ročný výnos
	zisk	ročný výnos	zisk	ročný výnos		
AUD	7.0123	0.6932	-287.1214	-28.4364	-166.3702	-16.4638
USD	181.3135	17.907	-122.6901	-12.1684	165.1587	16.3129

Tabuľka 25: Zisk/strata získané naivnou predikciou výmenných kurzov a stratégiou Buy & Hold

Ak porovnáme naivnú predikciu USD s výsledkami získanými ANN, zistíme, že priemerný výnos získaný zo simulácie obchodnej stratégie 2 použitím Levenberg-Marquardtovej metódy je vyšší ako výnos z naivnej predikcie. Hovoríme pritom o predpovedaní samotných výmenných kurzov. V prípade AUD je sieť v priemere horšia. Medzi 500 generovanými simuláciami sa však nájdu aj ziskovejšie, ako je naivná predikcia.

Pri porovnávaní stratégie Buy & Hold s ANN sme prišli k nasledovnému záveru. Ak nebudeme uvažovať transakčné náklady, sme schopní získať pomocou ANN v priemere

lepšiu predikciu. Ak však zahrnieme do simulácie obchodnej stratégie aj náklady, dostaneme pomocou ANN v priemere horšie výsledky.



Obrázok 13: Vývoj hodnoty investície získanej naivnou predikciou a stratégiou Buy & Hold

Záver

Cieľom tejto diplomovej práce bolo zhodnotiť možnosti využitia neurónových sietí na predpovedanie výmenných kurzov. Zamerali sme sa na technickú analýzu, kedy sme ako vstup do siete zadávali iba historické hodnoty daného časového radu. Predpovedali sme nielen samotné výmenné kurzy ale aj ich výnosy. Vychádzali sme pritom z výmenných kurzov eura voči americkému a austrálskemu doláru. Porovnali sme predikčnú schopnosť klasických neurónových sietí a neurónových sietí vyššieho rádu. Ukázalo sa, že v niektorých prípadoch sú lepšie HONN, v iných zase ANN v závislosti od modelovaného časového radu.

Pri predpovedaní kurzu eura voči americkému doláru, ktorý mal vyššie hodnoty Hurstovho koeficientu, sme dosiahli pomerne dobré výsledky. I priemerná výška zisku dosiahnutá simuláciou obchodnej stratégie 2 prekonalala výnos z naivnej predikcie. Stratégiu Buy&Hold sme však prekonalí iba bez zahrnutia transakčných nákladov. Výmenný kurz eura voči austrálskemu doláru dosahoval nižšie hodnoty Hurstovho koeficientu. To sa odrazilo v horších výsledkoch predpovedí.

Intuitívne sme prišli k záveru, že inicializácia váh je kľúčovým momentom pri predpovedaní. Avšak ani použitie váh, ktoré dávali najlepšie výsledky na trénovacej množine nevedlo ku tým najlepším predpovediam. Tento fenomén dokonale ilustruje hlavný problém vyhodnocovania dynamických systémov na dlhodobé obdobie. Nie je to problém len neurónových sietí. Špecifikum sietí spočíva v tom, že pre rôzne počiatkové váhy sa simuláciou dopracujeme ku rôznym váham a teda rôznym výsledkom predpovede. Preto by sme potrebovali kritérium, ktoré by vedelo čo najlepšie zhodnotiť nasimulovanú sieť. Ukázalo sa, že klasická MSE nie je smerodajným indikátorom dobrých výsledkov predpovede. Lepším kritériom sa ukázali byť D_{stat} a S_{stat} . Otázne však je, pri akých hodnotách D_{stat} resp. S_{stat} môžeme systém prehlásiť za ziskový. Ak máme systém, ktorý dokáže predpovedať smer výmenného kurzu s presnosťou 60%, ešte nemusíme byť schopný vytvoriť zisk. Podstatná je skôr veľkosť zmien, ktoré sú správne predpovedané. Ak systém správne odhadne transakcie s nízkymi výnosmi a nesprávne transakcie s vysokými výnosmi, stále môže produkovať stratu. Tento vzťah nie je zahrnutý v D_{stat} ani v S_{stat} , preto by to možno chcelo nejaké efektívnejšie kritérium, na základe ktorého sa rozhodovať o výkonnosti danej siete.

Ako sme videli, výnos je do značnej miery podmienený použitím obchodnej stratégie, teda tým, ako sme schopní využiť predpovedanú informáciu. Ukázalo sa, že stratégia 2 je výhodnejšia v prípade amerického dolára. V prípade austrálskeho dolára je výhodnejšia naopak stratégia 1. Pri oboch stratégiách vidíme, že efekty transakčných nákladov sú pomerne vysoké. V praxi sa však môžu použiť aj iné obchodné stratégie, v závislosti od investora a od jeho postoju k riziku.

Jedným s cieľov práce bolo porovnať schopnosti neurónových sietí a klasických štatistických metód. Tu sme však narazili na obmedzenia vyplývajúce z množstva predpokladov, ktoré by mali byť splnené. Hlavný obmedzujúci predpoklad ARMA modelov je v tom, že očakáva konštantnú volatilitu. Táto vlastnosť však pri dátach, akými sú výnosy finančných časových radov, nie je splnená. Preto sme sa rozhodli aplikovať klasické metódy na predpovedanie volatility výnosov, na čo sme využívali modely GARCH resp. ARMA+GARCH.

Poznatky získané v práci potvrdili, že neurónové siete majú potenciál na predpovedanie finančných časových radov. Avšak aj my sme čelili nedostatkom, ktoré sú spojené s ich využitím. Pri neurónových sieťach je dôležitá voľba architektúry a parametrov siete. To si však vyžaduje istú dávku skúseností, pretože voľba závisí od konkrétneho problému. Pomerne často sa aplikuje metóda pokusov a omylov. Táto procedúra je však veľmi zdĺhavá. Na nájdenie uspokojivej architektúry je potrebné opakované tréningovanie modelu, pričom výber nie je vždy jednoznačný. Otázna je hlavne voľba kritéria, na základe ktorého určíme tú najlepšiu sieť. Pri sieťach sa stretávame s problémom interpretácie výsledkov. V protiklade s tradičnými metódami, ktoré majú len malý počet parametrov, výstupy z neurónových sietí sa interpretujú pomerne zložito práve kvôli ich vysokému počtu. Táto vlastnosť je limitujúcim faktorom nielen pri použití neurónových sietí na predpovedanie finančných časových radov, ale je to problém neurónových sietí ako takých.

Dospeli sme k záveru, že neurónové siete by mohli za predpokladu správneho tréningovania a správnej voľby architektúry úspešne predpovedať budúci vývoj. I keď preukázali vysoký potenciál, je len málo pravdepodobné, že ich predikčné schopnosti výrazne porastú. Veľké dynamické systémy ako finančné trhy sú jednoducho príliš komplikované na to, aby mohli byť predpovedané na dlhý čas. Použitie neurónových sietí na predpovedanie bude aj v budúcnosti v centre záujmu, keďže investori hľadajú akékoľvek nástroje na zlepšovanie svojich predpovedí a prekonanie trhu s jediným cieľom. Zvýšiť svoj zisk.

Možné zlepšenia

Neurónové siete sme v tejto práci testovali len na dvoch výmenných kurzoch. Pre objektívne zhodnotenie výsledkov by bolo potrebné preskúmať väčší počet radov. Navyše sme sa v tejto práci zaoberali iba predpovedaním na jeden krok dopredu. Z ekonomického hľadiska by bolo možno zaujímavejšie vyčíslovať predpovede na viacej dní dopredu.

Hlavné zameranie tejto práce bola technická analýza. Samozrejme, použitie iba historických dát daného časového radu bez žiadnych iných vysvetľujúcich premenných je pomerne obmedzujúce, ale samo o sebe tiež veľmi zaujímavé. Pri fundamentálnej analýze vychádzame zo závislostí medzi jednotlivými časovými radmi. Zaujímavé by bolo zahrnutie ďalších výmenných kurzov, akciových indexov, úrokovej miery alebo miery nezamestnanosti. To by mohlo do značnej miery vylepšiť výsledok získaný z predpovedí.

V tejto práci sme pracovali s doprednými neurónovými sieťami a neurónovými sieťami druhého rádu. Na predpovedanie sa však používajú aj iné architektúry. Každá z architektúr má svoje výhody a nevýhody. Dopredné neurónové siete tréningované metódou spätného šírenia

chyby sú bežné, i keď ich konfigurácia a tréovanie je často zložitú. Použili tzv. okno do minulosti, ktorým sme vyjadrili časovú štruktúru skrytú v dátach. Reprezentácie časového kontextu pomocou okna do minulosti konečnej nemennej dĺžky nemusí byť vždy dostatočne silným nástrojom na zvládnutie časovej štruktúry dát. Oproti tomu rekurentné neurónové siete majú vlastnú pamäť, čo im umožňuje extrahovať z dát časové závislosti. Táto ich vlastnosť môže zlepšovať predikciu. Preto by bolo zaujímavé porovnať výsledky získané doprednými a rekurentnými neurónovými sieťami. Okrem nich existujú aj iné komplikovanejšie modely, ktoré môžu znižovať chybu predpovede. Na druhej strane sú však často ťažšie tréovateľné a analyzovateľné.

Literatúra

- [1] AMBRIŠKO R., *Aplikácia teórie chaosu na menovom trhu Slovenskej Republiky*, Diplomová práca, Ekonomická univerzita, Podnikovohospodárska fakulta, 2003
- [2] CONNOR J. T., MARTIN R. D., ATLAS L. E., *Recurrent neural networks and robust time series prediction*, IEEE Transactions on neural networks, vol 5, no. 2, March 1994
- [3] DEMUTH H., BEALE M., HAGAN M., *Matlab neural network toolbox 5*, User's guide, MathWorks, 2007
- [4] DERENÍK D., *Matematické modely v technickej analýze cenných papierov*, Bakalárska práca, Masarykova univerzita, Přírodovědecká fakulta, 2006
- [5] *Econometrics toolbox 1*, User's guide, MathWorks, 2008
- [6] GIACOMINI E., *Neural networks in quantitative finance*, Master Thesis, Humboldt Universität, Center for Applied Statistics and Economics, 2003
- [7] GRADOJEVIC N., YANG J., *The application of artificial neural networks to exchange rate forecasting: The role of market microstructure variables*, Bank of Canada, Working Paper 2000-23
- [8] HAYKIN S., *Neural networks: A comprehensive foundation (2nd ed.)*, Prentice Hall, NJ, USA, 1999
- [9] HERČÚT R., *Predpovedanie výmenného kurzu pomocou neurónových sietí*, Diplomová práca, Univerzita Mateja Bela, Fakulta prírodných vied, 2008
- [10] KUNST R., *Econometric forecasting*, Lecture Notes, University of Vienna, Department of Economics, February 2009
- [11] KVASNIČKA V. a spol, *Úvod do teórie neurónových sietí*, Iris, Bratislava, 1997
- [12] LATOURETTE K., *A comparison of the Levenberg-Marquardt method with standard optimization algorithm, in minimizing the Tikhonov-total variational functional*, University of Arizona, Applied Mathematics, 2008
- [13] LE BOUDEC J.-Y., *Performance evaluation of computer and communication system*, Lecture notes, École Polytechnique Fédérale de Lausanne, Confédération Helvétique
- [14] LI C., *Rescaled-range and power spectrum analyses on well-logging data*, Geophysical Journal International, (2003) 153, pp 201–212

- [15] LJUNG L., *System identification toolbox 7*, User's guide, MathWorks, 2008
- [16] MCNELIS P. D., *Neural networks in finance, Gaining predictive edge in the market*, Elsevier Academic Press, USA, 2005
- [17] NÁVRAT P. a kol., *Umelá Inteligencia*, STU, Bratislava, 2002
- [18] NYGREN K., *Stock prediction – a neural network approach*, Master Thesis, Royal Institute of Technology, KTH, 2004
- [19] *Optimization toolbox 4*, User's guide, MathWorks, 2008
- [20] PALLIKARI F., BOLLER E., *A rescaled range analysis of random events*, Journal of Scientific Exploration, Vol. 13, No. 1, pp. 25–40, 1999
- [21] PANG J., NORTH C.P., *Fractals and their applicability in geological wireline log analysis*, Journal of Petroleum Geology, vol 19(3), 1996, pp 339-350
- [22] PISSARENKO D., *Neural networks for financial time series prediction: Overview over recent research*, 2002
- [23] QIAN B., RASHEED K., *Hurst exponent and financial market predictability*, Financial Engineering and Applications, 437-043, 2004
- [24] ROWEIS S., *Levenberg-Marquardt optimization*, Notes, University Of Toronto
- [25] SCHWAERZEL R., *Improving the prediction accuracy of financial time series by using multi-neural network systems and enhanced data preprocessing*, Thesis, University of Texas, 1996
- [26] SINČÁK P., ANDREJKOVÁ G., *Neurónové siete I. (Inžiniersky prístup)*, ELFA Press, Košice, 1996
- [27] SINČÁK P., ANDREJKOVÁ G., *Neurónové siete II. (Inžiniersky prístup)*, ELFA Press, Košice, 1996
- [28] WÜRTZ D., CHALABI Y., LUKSAN L., *Parameter estimation of ARMA models with GARCH/APARCH errors, An R and SPlus software implementation*, Journal of Statistical Software, Draft
- [29] XU S., CHEN L., *A novel approach for determining the optimal number of hidden layer neurons for FNN's and its application in data mining*, 5th International Conference on Information Technology and Applications (ICITA 2008), June 2008, Cairns, Queensland, Australia.
- [30] YAO J. T., POH H.-L., *Equity forecasting: A case study on the KLSE index*, Proceedings of the 1995 IEEE International Conference on Neural Networks (ICNN'95), Perth, Western Australia, Nov 1995, Volume 2, pp1012-1017

- [31] YAO J. T., TAN C. L., *A case study on using neural networks to perform technical forecasting of forex*, Neurocomputing 34 (2000) 79-98, Elsevier Science
- [32] YAO J. T., TAN C. L., *A study on training criteria for financial time series forecasting*, Proceedings of International Conference on Neural Information Processing, Shanghai, China, November, 2001, pp772-777
- [33] YU L., WANG S., LAI K. K., *A novel nonlinear ensemble forecasting model incorporating GLAR and ANN for foreign exchange rates*, Computers and Operations Research, Volume 32, Issue 10 (2005), pp 2523 – 2541
- [34] YU L., WANG S., LAI K. K., *Adaptive smoothing neural networks in foreign exchange rate forecasting*, Springer Berlin / Heidelberg, Volume 3516/2005, Pages 523-530
- [35] ZHANG M., *Artificial higher order neural networks for economics and business*, Information science reference, Hershey, New York, USA, 2009

Príloha A

Hurstov koeficient a R/S analýza

Intuitívna fraktálna povaha finančných časových radov viedla ku aplikovaniu teórie fraktálov a chaosu pri ich analýze. Hurstov koeficient predstavuje štatistickú mieru používanú na klasifikáciu časových radov. Keďže ide o priamočiary nástroj vyžadujúci iba málo predpokladov, má rozsiahle použitie pri analýze časových radov. V praxi sa stretávame s viacerými metódami, ktoré sa používajú na odhad Hurstovho koeficientu. Presnosť odhadov je však komplikovaným problémom. Jednou z najčastejšie používaných metód je škálovanie rozsahu alebo R/S analýza.

Majme časový rad $X_1, X_2 \dots X_n$, z ktorého vyberieme vzorku n dát. R/S analýza prebieha nasledovne:

- vytvoríme akumulované odchýlky od priemernej hodnoty

$$Y_{t,n} = \sum_{i=1}^t (X_i - \bar{X}), \quad t = 1, 2, \dots, n,$$

kde \bar{X} je priemer z hodnôt X_1 až X_n .

- vypočítame rozsah určený ako rozdiel maximálnej a minimálnej akumulovanej výchyľky za dané obdobie

$$R_n = \max_{1 \leq t \leq n} (Y_{t,n}) - \min_{1 \leq t \leq n} (Y_{t,n})$$

- vypočítame štandardné odchýlky

$$S_n = \sqrt{\frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2}$$

- rozsah normalizujeme štandardnou odchýlkou, čím dostaneme škálovací rozsah $(R/S)_n$

$$(R/S)_n = R_n / S_n$$

Veľkosť (R/S) rastie s časom podľa nasledovnej rovnice

$$(R/S)_n \approx c n^H,$$

kde c je konštanta a H je Hurstov koeficient. Existuje viacero metód, ako odhadnúť hodnotu Hurstovho koeficientu. Jednou z možností je výpočet sklonu regresnej priamky, ktorá reprezentuje graf $\log(R/S)$ vs. $\log(n)$. Tento vzťah môžeme analyticky vyjadriť rovnicou

$$\log(R/S)_n = \log c + H \log n .$$

Na zabezpečenie presnosti odhadu potrebujeme aspoň 10 hodnôt. Hodnoty R/S použité v regresii sú priemerné hodnoty pre každé n . Ak máme celkový objem dát n_t , priemerné hodnoty R/S budú odhadnuté nasledovne: Množina n_t bude rozdelená do v skupín, každá skupina pri tom bude obsahovať $n = n_t/v$ dát. Parameter v nadobúda také hodnoty od 1 do n_t , aby n bolo prirodzené číslo a $n \geq 4$. R/S je odhadnuté pre každú z v skupín a ich priemer $(R/S)_{\text{avg}}$ je potom zobrazený v log-log grafe voči n .

Interpretácia Hurstovho koeficientu

Hodnoty Hurstovho koeficientu sú z rozmedzia 0 až 1. Na ich základe klasifikujeme rady do troch kategórií:

- $H=0.5$ indikuje náhodný časový rad
- $0 \leq H < 0.5$ indikuje antiperzistentný časový rad
- $0.5 < H \leq 1$ indikuje perzistentný časový rad

$H=0.5$ naznačuje, že ide o náhodný proces pozostávajúci z nekorelovaných udalostí. Keďže ide o neparametrickú metódu, ktorá nekladie podmienky na pravdepodobnostné rozdelenie, môže byť rozdelenie náhodného procesu ľubovoľné. V prípade, že $H \neq 0.5$, pozorovania nie sú nezávislé.

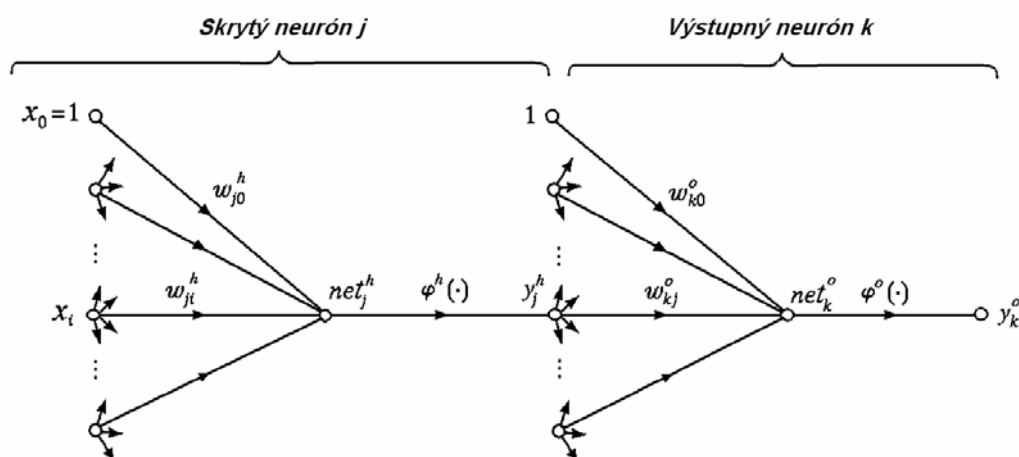
Antiperzistentný časový rad sa vyznačuje častými zmenami a vysokou volatilitou. Jeho hodnoty sa vracajú ku svojej strednej hodnote. To znamená, že vysoká hodnota je s veľkou pravdepodobnosťou nasledovaná nízkou hodnotou a naopak. Sila tejto závislosti rastie s poklesom hodnoty Hurstovho koeficientu. Napriek tomu, že je tento koncept v ekonomickej a finančnej literatúre rozšírený, doteraz bolo identifikovaných len málo antiperzistentných radov [31].

Väčšina ekonomických časových radov je teda perzistentná s $H > 0.5$. Ide o rady obsahujúce trend, ktoré sú charakterizované efektom dlhodobej pamäte. Sila tejto závislosti rastie s nárastom Hurstovho koeficientu. Čím sú hodnoty H menšie, tým je v systéme viacej šumu a časový rad sa viac ponáša na náhodný pohyb. Naopak, čím je H väčšie, tým obsahuje menej šumu.

Príloha B

Odvedenie metódy spätného šírenia chyby

V tejto prílohe popíšeme postup odvedenia metódy spätného šírenia chyby pre dvojvrstvovú doprednú neurónovú sieť. Predpokladajme, že máme m vstupných neurónov, n neurónov v skrytej vrstve a N výstupných neurónov. Indexom o budeme označovať veličiny vzťahujúce sa k výstupnej vrstve, zatiaľ čo index h bude použitý na označenie skrytej vrstvy. Váhy medzi vstupnou a skrytou vrstvou budeme reprezentovať symbolom w_{ji}^h , zatiaľ čo symbol w_{kj}^o bude použitý na označenie spojenia medzi skrytou a výstupnou vrstvou. Podrobnejší popis vidíme na obrázku.



Neuróny vo výstupnej vrstve

Uvažujme sieť znázornenú na obrázku vyššie. Výstupné neuróny spracovávajú signály vytvárané predchádzajúcou skrytou vrstvou. Výsledný signál vytváraný k -tým výstupným neurónom je teda

$$y_k^o = \varphi^o(\text{net}_k^o) = \varphi^o\left(\sum_{j=0}^n w_{kj}^o y_j^h\right).$$

Chybový signál výstupného neurónu je definovaný ako rozdiel medzi požadovanou a skutočnou hodnotou výstupu. Celková chyba siete je určená súčtom chýb vo všetkých výstupných neurónoch

$$E = \frac{1}{2} \sum_{k=1}^N e_k^2 = \frac{1}{2} \sum_{k=1}^N (d_k - y_k^o)^2 .$$

Dosadením za y_k^o dostaneme

$$E = \frac{1}{2} \sum_{k=1}^N \left(d_k - \varphi^o \left(\sum_{j=0}^n w_{kj}^o y_j^h \right) \right)^2 .$$

Metóda spätného šírenia chyby podobne ako delta pravidlo upravuje veľkosť váh v závislosti od výšky záporného gradientu chybovej funkcie, teda

$$w_{kj}^o(t+1) = w_{kj}^o(t) - \eta \frac{\partial E}{\partial w_{kj}^o} .$$

Veľkosť tejto parciálnej derivácie môžeme vyjadriť vzt'ahom

$$\frac{\partial E}{\partial w_{kj}^o} = \frac{\partial E}{\partial e_k} \frac{\partial e_k}{\partial y_k^o} \frac{\partial y_k^o}{\partial net_k^o} \frac{\partial net_k^o}{\partial w_{kj}^o} ,$$

pričom

$$\begin{aligned} \frac{\partial E}{\partial e_k} &= e_k = d_k - y_k^o & \frac{\partial y_k^o}{\partial net_k^o} &= \frac{\partial \varphi^o}{\partial net_k^o} = \varphi^o{}'(net_k^o) \\ \frac{\partial e_k}{\partial y_k^o} &= -1 & \frac{\partial net_k^o}{\partial w_{kj}^o} &= y_j^h . \end{aligned}$$

Dosadením do gradientu chybovej funkcie dostávame

$$\frac{\partial E}{\partial w_{kj}^o} = -(d_k - y_k^o) \varphi^o{}'(net_k^o) y_j^h .$$

Na základe predošlých výpočtov budú váhy vo výstupnej vrstve modifikované vzorcom

$$w_{kj}^o(t+1) = w_{kj}^o(t) + \eta (d_k - y_k^o) \varphi^o{}'(net_k^o) y_j^h .$$

V odbornej literatúre sa často stretávame s výrazom

$$\Delta w_{kj}^o = \eta \delta_k^o y_j^h ,$$

kde

$$\begin{aligned} \delta_k^o &= - \frac{\partial E}{\partial e_k} \frac{\partial e_k}{\partial y_k^o} \frac{\partial y_k^o}{\partial net_k^o} \\ &= (d_k - y_k^o) \varphi^o{}'(net_k^o) . \end{aligned}$$

V praktických úlohách sa často používajú výstupné neuróny s lineárnou aktivačnou funkciou. Pre lineárny výstupný neurón dostaneme vzt'ah

$$\Delta w_{kj}^o = \eta e_k y_j^h .$$

Neuróny v skrytej vrstve

V prípade skrytých neurónov je situácia trochu zložitejšia. Ako vidíme z rovnice pre zmenu váh výstupných neurónov, kľúčovým faktorom je výška chybového signálu. V prípade výstupných neurónov, ktorým je priamo priradená požadovaná hodnota výstupu, je výpočet chybového signálu priamočiary. Skryté neuróny však nemajú priradenú žiadnu požadovanú hodnotu, pretože nie sú priamo dostupné. Napriek tomu sa aj tieto neuróny podieľajú na výslednej chybe celej siete. Výška chybového signálu je určená rekurzívne pomocou chybových signálov neurónov, s ktorými je priamo spojený. Hovoríme o spätnom šírení chybového signálu.

Neuróny v skrytej vrstve spracovávajú priamo vstupný signál, teda hodnota výstupu zo skrytého neurónu je

$$y_j^h = \varphi^h(\text{net}_j^h) = \varphi^h\left(\sum_{i=0}^m w_{ji}^h x_i\right).$$

Chyba siete je definovaná stále rovnako, teda

$$E = \frac{1}{2} \sum_{k=1}^N e_k^2 = \frac{1}{2} \sum_{k=1}^N (d_k - y_k^o)^2.$$

Dosadením za y_k^o a y_j^h dostaneme

$$E = \frac{1}{2} \sum_{k=1}^N \left(d_k - \varphi^o \left(\sum_{j=0}^n w_{kj}^o \varphi^h \left(\sum_{i=0}^m w_{ji}^h x_i \right) \right) \right)^2.$$

Váhy skrytých neurónov sú modifikované rovnakým spôsobom ako v prípade výstupných neurónov

$$w_{ji}^h(t+1) = w_{ji}^h(t) - \eta \frac{\partial E}{\partial w_{ji}^h}.$$

Gradient chybovej funkcie vypočítame nasledovným spôsobom

$$\frac{\partial E}{\partial w_{ji}^h} = \frac{\partial E}{\partial y_j^h} \frac{\partial y_j^h}{\partial \text{net}_j^h} \frac{\partial \text{net}_j^h}{\partial w_{ji}^h},$$

pričom

$$\begin{aligned} \frac{\partial E}{\partial y_j^h} &= \sum_{k=1}^N e_k \frac{\partial e_k}{\partial y_j^h} \\ &= -\sum_{k=1}^N (d_k - y_k^o) \frac{\partial \varphi^o(\text{net}_k^o)}{\partial y_j^h} & \frac{\partial y_j^h}{\partial \text{net}_j^h} &= \frac{\partial \varphi^h}{\partial \text{net}_j^h} = \varphi^h'(\text{net}_j^h) \\ &= -\sum_{k=1}^N (d_k - y_k^o) \frac{\partial \varphi^o(\text{net}_k^o)}{\partial \text{net}_k^o} \frac{\partial \text{net}_k^o}{y_j^h} & \frac{\partial \text{net}_j^h}{\partial w_{ji}^h} &= x_i. \\ &= -\sum_{k=1}^N (d_k - y_k^o) \varphi^o'(\text{net}_k^o) w_{kj}^o \end{aligned}$$

Dosadením do gradientu chybovej funkcie dostávame

$$\frac{\partial E}{\partial w_{ji}^h} = -\sum_{k=1}^N (d_k - y_k^o) \rho'(\text{net}_k^o) w_{kj}^o \varphi'(\text{net}_j^h) x_i.$$

Váhy skrytých neurónov budú modifikované vzorcom

$$w_{ji}^h(t+1) = w_{ji}^h(t) + \eta \sum_{k=1}^N (d_k - y_k^o) \rho'(\text{net}_k^o) w_{kj}^o \varphi'(\text{net}_j^h) x_i.$$

V odbornej literatúre má tento výraz tvar

$$\Delta w_{ji}^h = \eta \delta_j^h x_i,$$

kde

$$\delta_j^h = \varphi'(\text{net}_j^h) \sum_{k=1}^N \delta_k^o w_{kj}^o.$$

Príloha C

Predpovedanie volatility

Neurónové siete sa môžu používať nielen na predpovedanie aktív a ich výnosov, ale aj na predpovedanie podmienenej volatility. Na modelovanie volatility sa zvyčajne používajú ARCH resp. GARCH modely, prípadne ich modifikácie. Neurónové siete budeme aplikovať na ARCH(q) model, ktorý môžeme zapísať v tvare

$$y_t = c + \varepsilon_t \quad \varepsilon_t \sim D(0, \sigma_t^2)$$
$$\sigma_t^2 = \kappa + \alpha_1 \varepsilon_{t-1}^2 + \alpha_2 \varepsilon_{t-2}^2 + \dots + \alpha_q \varepsilon_{t-q}^2.$$

Pre naše účely bude lepší nasledovný zápis

$$y_t = c + \psi(y_{t-1}, y_{t-2}, \dots, y_{t-q}) \varepsilon_t \quad \varepsilon_t \sim D(0, 1),$$

resp.

$$y_t = c + \psi(z_t) \varepsilon_t \quad \varepsilon_t \sim D(0, 1),$$

kde $z_t = (y_{t-1}, y_{t-2}, \dots, y_{t-q})^T \in \mathfrak{R}^q$. Výšku podmienenej strednej hodnoty a disperzie výnosov vypočítame z predošlej rovnice ako

$$E(y_t | z_t = z) = c$$

$$D(y_t | z_t = z) = \psi^2(z).$$

Na odhad strednej hodnoty výnosov môžeme použiť aritmetický priemer y_t , ktorý je rovnomerne najlepším lineárne nevychýleným odhadom strednej hodnoty. Označíme ho \hat{c} . Naším cieľom však bude určiť podmienenú volatilitu, teda veľkosť $\hat{\psi}^2(z)$. Na to využijeme rezíduá ε_t , ktoré môžeme zapísať v tvare

$$\varepsilon_t = y_t - c = \psi(z_t) \varepsilon_t.$$

Majú nasledovné vlastnosti

$$E(\varepsilon_t | z_t = z) = 0$$

$$D(\varepsilon_t | z_t = z) = \psi^2(z) \quad \Rightarrow \quad E(\varepsilon_t^2 | z_t = z) = \psi^2(z).$$

Aproximáciou rezíduí $\hat{\varepsilon}_t = y_t - \hat{c}$ a následnou simuláciou $\hat{\varepsilon}_t^2$ pomocou neurónových sietí získame odhad $\hat{\psi}^2(z)$. Z neho už ľahko získame výšku podmienenej volatility.