An Increasing Hybrid Morphological-Linear Perceptron with Evolutionary Learning and Phase Correction for Financial Time Series Forecasting

Ricardo de A. Araújo Peter Sussner

Information Technology Department, [gm]^2 Intelligent Systems, Brazil
Department of Applied Mathematics, University of Campinas, Brazil

5th International Conference on Hybrid Artificial Intelligence Systems

Organization of this talk

Organization of this talk

1. Introduction
2. Morphological Neural Networks
3. The Increasing Hybrid Morphological-Linear Perceptron
4. Experimental Results
5. Concluding Remarks

Organization of this talk

1. Introduction
2. Morphological Neural Networks
3. The Increasing Hybrid Morphological-Linear Perceptron
4. Experimental Results
5. Concluding Remarks

Organization of this talk

Organization of this talk

Morphological Neural Networks

We speak of a morphological neural network (MNN) if every neuron performs an elementary operation of mathematical morphology (MM).

MNNs are closely related to other lattice-based neurocomputing models.

This talk presents a hybrid morphological/linear neural network for financial time series prediction.

Observations on Financial Time Series

1. Financial time series (FTS) exhibit a strong random walk linear component and a weaker nonlinear component;
2. Experimental results indicate that FTS can be modelled as increasing functions (from the domain of time lags).

The Increasing Hybrid Morphological-Linear Perceptron (IHMP)

Based on these observations, we propose a hybrid model, called increasing hybrid morphological-linear perceptron (IHMP), consisting of a convex combination of

1. a conventional perceptron (linear part);
2. an increasing morphological perceptron (nonlinear part).

The learning process of the proposed IHMP includes an automatic phase correction step.

Basic Concepts of Morfological Neural Networks

Context of this Talk

- Mathematical morphology (MM) is concerned with the *processing and analysis of images using structuring elements*;
- Complete lattices provide for the appropriate algebraic framework of MM;
- The elementary operations of mathematical morphology can be defined in this complete lattice framework;
- The neurons of a morfological neural network (MNN) perform elementary operations of mathematical morphology, possibly followed by an activation function.
- Applications of MNNs include *classification, character recognition, automatic target recognition (in particular landmine detection), image reconstruction, image compression, and time serie prediction*.

Some Pertinent Notions of Lattice Theory

A complete lattice is a partially ordered set $\mathbb{L}$ such that every $Y \subseteq \mathbb{L}$ has an infimum, denoted by $\bigwedge Y$ and a supremum, denoted by $\bigvee Y$ in $\mathbb{L}$.

From now on, the symbols $\mathbb{L}$ and $\mathbb{M}$ denote complete lattices. $\mathbb{L}^n$ is also a complete lattice with the partial order given by

$$(x_1, \ldots, x_n) \leq (y_1, \ldots, y_n) \Leftrightarrow x_i \leq y_i, \; i = 1, \ldots, n$$

Examples of complete lattices include $\mathbb{R}_{\pm\infty} = \mathbb{R} \cup \{+\infty, -\infty\}$, $\mathbb{R}^n_{\pm\infty} = (\mathbb{R}_{\pm\infty})^n$, $[0, 1]$, and $[0, 1]^{\mathbf{X}}$.

Some Basic Operators of MM on Complete Lattices

Erosion

An operator $\varepsilon : \mathbb{L} \to \mathbb{M}$ represents an (algebraic) erosion if

$$\varepsilon\left(\bigwedge Y\right) = \bigwedge_{y \in Y} \varepsilon(y), \quad \forall Y \subseteq \mathbb{L}.$$

Dilation

An operator $\delta : \mathbb{L} \to \mathbb{M}$ represents a (algebraic) dilation if

$$\delta\left(\bigvee Y\right) = \bigvee_{y \in Y} \delta(y), \quad \forall Y \subseteq \mathbb{L}.$$

Specific Examples of Erosion and Dilation

Max Product and Min Product

For $A \in \mathbb{R}^{m \times p}$ and $B \in \mathbb{R}^{p \times n}_{\pm\infty}$, the max-product $C = A \boxdot\!\!\vee B$ and the min-product $D = A \boxdot\!\!\wedge B$ are defined by

$$c_{ij} = \bigvee_{k=1}^{p} (a_{ik} + b_{kj}), \quad d_{ij} = \bigwedge_{k=1}^{p} (a_{ik} + b_{kj}) \,.$$

For $A \in \mathbb{R}^{n \times m}$, the following operators $\varepsilon_A, \delta_A : \mathbb{R}^n_{\pm\infty} \to \mathbb{R}^m_{\pm\infty}$ represent respectively an (algebraic) erosion and dilation.

$$\varepsilon_A(\mathbf{x}) = A^T \boxdot\!\!\wedge \mathbf{x} \,, \ \delta_A(\mathbf{x}) = A^T \boxdot\!\!\vee \mathbf{x} \,.$$

In the near future, we intend to prove that all erosions and dilations $\mathbb{R}^n_{\pm\infty} \to \mathbb{R}^m_{\pm\infty}$ are of this form.

Decomposition of Increasing Mappings

A mapping $\Psi : \mathbb{L} \to \mathbb{M}$ is called increasing if

$$x \leq y \Rightarrow \Psi(x) \leq \Psi(y) \; \forall x, y \in \mathbb{L}.$$

Banon and Barrera Decompositions (B & B)

Let $\Psi : \mathbb{L} \to \mathbb{M}$ be increasing. There exist erosions ε^i and dilations δ^j for some index sets I and J such that

$$\Psi = \bigvee_{i \in I} \varepsilon^i = \bigwedge_{j \in J} \delta^j .$$

For increasing $\Psi : \mathbb{R}^n \to \mathbb{R}$, our hypothesis and B & B suggest that there exist $\mathbf{v}^i, \mathbf{w}^j \in \mathbb{R}^n$ and finite I^*, J^* such that

$$\Psi \simeq \bigvee_{i \in I^*} \varepsilon_{\mathbf{v}^i} \quad \text{and} \quad \Psi \simeq \bigwedge_{j \in J^*} \delta_{\mathbf{w}^j} .$$

The Proposed IHMP Models

Motivation

Experiments indicate that the FTS we considered are given by increasing functions $\Psi : \mathbb{R}^n \to \mathbb{R}$, where n represents the number of antecedents or time lags.

Definition of our IHMP Models

Given input $\mathbf{x} \in \mathbb{R}^n$, the following IHMPs calculate

$$y = \lambda\alpha + (1-\lambda)\beta, \qquad \lambda \in [0,1]\,,$$

where α represents the increasing morphological module and

$$\beta = \mathbf{x} \cdot \mathbf{b}^T = x_1 b_1 + x_2 b_2 + \ldots + x_n b_n$$

represents the linear module.

Erosion-Based and Dilation-Based IHMPs

Erosion-Based IHMP (E-IHMP)

$$\alpha = \bigvee_{i=1}^{k} v_i \quad \text{where} \quad v_i = \varepsilon_{\mathbf{a}^i}(\mathbf{x}) = \bigwedge_{j=1}^{n} (a_j^i + x_j) \, .$$

Dilation-Based IHMP (D-IHMP)

$$\alpha = \bigwedge_{i=1}^{k} v_i \, , \quad \text{where} \quad v_i = \delta_{\mathbf{a}^i}(\mathbf{x}) = \bigvee_{j=1}^{n} (a_j^i + x_j) \, .$$

In both models $\mathbf{a}^i = (a_1^i, a_2^i, \ldots, a_n^i)^T \in \mathbb{R}^n$.

Architectures of IHMPs

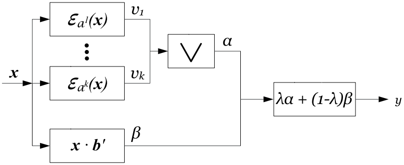

Figure: Architecture of E-IHMP.

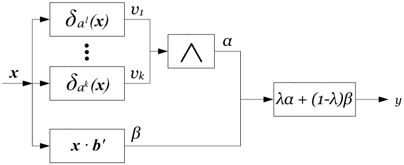

Figure: Architecture of D-IHMP.

Evolutionary Training Algorithm for IHMP Models

If $\mathbf{a}^T = ((\mathbf{a}^1)^T, (\mathbf{a}^2)^T, \ldots, (\mathbf{a}^k)^T)$ then the weight vector $\mathbf{w}$ of both the E-IHMP and the D-IHMP is given by

$$\mathbf{w}^T = (\lambda, \mathbf{a}^T, \mathbf{b}^T)\,.$$

Let $d(m)$ and $y(m)$ be respectively the desired output and the actual output for the m-th training pattern, where $m = 1, \ldots, M$. Define the following fitness function $f(\mathbf{w})$:

$$f(\mathbf{w}) = \frac{1}{1 + \sum_{m=1}^{M} e^2(m)} \quad \text{where} \quad e(m) = d(m) - y(m)\,.$$

Initialization and Stopping Criteria

Initialization

- Random initialization of **a** and **b** within the range $[-1, 1]$;
- Random initialization of λ within the range $[0, 1]$;
- Choice of k varies for each prediction problem.

Stopping Criteria

- Maximum generation number $gen = 10000$;
- Training error $Pt \leq 10^{-6}$;
- Increase of the validation error or generalization loss of the fitness function $> 5\%$.

Modified Genetic Algorithm (MGA)

Use *roulette wheel* approach to obtain $\mathbf{p}^1$ and $\mathbf{p}^2$ and generate

$$\begin{aligned}
\mathbf{C}^1 &= \frac{\mathbf{p}^1 + \mathbf{p}^2}{2}\,, \\
\mathbf{C}^2 &= w(\mathbf{p}^1 \vee \mathbf{p}^2) + (1-w)\mathbf{p}^{max}\,, \\
\mathbf{C}^3 &= w(\mathbf{p}^1 \wedge \mathbf{p}^2) + (1-w)\mathbf{p}^{min}\,, \\
\mathbf{C}^4 &= \frac{w(\mathbf{p}^1 + \mathbf{p}^2) + (1-w)(\mathbf{p}^{max} + \mathbf{p}^{min})}{2}\,,
\end{aligned}$$

where $w \in [0,1]$ (here 0.9) and $\mathbf{p}^{max}$, $\mathbf{p}^{min}$ *have max., min. gene values*. If $\mathbf{C}^{best}$ is the son with the *highest fitness* value then

$$\mathbf{MC}^j = \mathbf{C}^{best} + \mathbf{B}^j \Delta \mathbf{M}^j,\; j = 1,2,3\,,$$

where $\mathbf{p}^{min} \leq \mathbf{C}^{best} + \Delta\mathbf{M}^j \leq \mathbf{p}_{max}$ and $\mathbf{B}^j$ are certain binary vectors. *Some* $\mathbf{MC}^j$ *are incorporated* into population.

Automatic Phase Correction

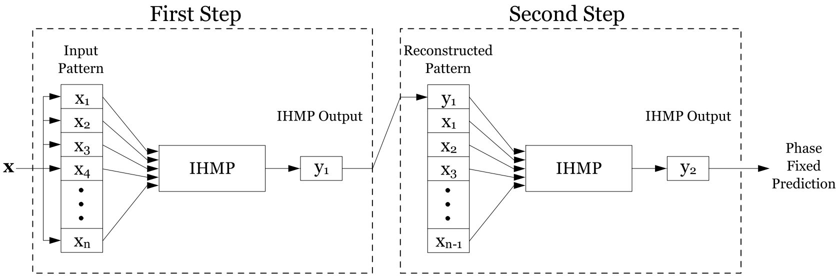

Figure: Phase fix procedure.

Experimental Results

Tests were performed using the

- *Dow Jones Industrial Average (DJIA) index*;
- *Standard & Poor 500 (S&P500) index*.

The data were normalized in the range $[0, 1]$ and divided into

- training set (50% of the data);
- validation set (25%);
- test set (25%).

We compared our IHMP models with

- ARIMA;
- multi-layer perceptron (MLP);
- modular morphological neural network (MMNN);
- morphological-rank-linear perceptron (MRL).

Performance Measures

- mean square error (MSE);
- mean absolute percentage error (MAPE);
- U of THEIL Statistics (THEIL);
- average relative variance (ARV);
- prediction of change in direction (POCID).

In addition, we employed the following evaluation function (EF):

$$\mathrm{EF} = \frac{\mathrm{POCID}}{1 + \mathrm{MSE} + \mathrm{MAPE} + \mathrm{THEIL} + \mathrm{ARV}}.$$

Results for the DJIA Test Set

- daily records $01/01/1998 - 08/26/2003$ (1420 points);
- input vectors comprise lags $2, 3, \ldots, 11$;
- $\#$ of basic morphological operations in IHMPs: $k = 8$.

Metrics	ARIMA	MMNN	MLP	MRL	**D-IHMP**	E-IHMP
MSE	5.8033e-4	8.3236e-4	8.3000e-2	8.2148e-4	**1.6044e-4**	1.7619e-4
MAPE	8.3200e-2	9.6700e-2	9.3788e-2	9.6578e-2	**5.7717e-2**	6.0262e-2
THEIL	1.2649	0.9945	0.9885	0.9916	**0.4965**	0.5094
ARV	3.9200e-2	3.4423e-2	3.4204e-2	3.3981e-2	**6.5683e-3**	7.2129e-3
POCID	46.10	50.85	46.59	46.82	**100.00**	**100.00**
EF	19.3058	23.9130	21.1822	22.0539	**64.0637**	63.4095

Results for the S&P500 Test Set

- monthly records $01/1970 - 08/2003$ (369 points);
- input vectors comprise lags $2, 3, \ldots, 6$;
- $\#$ of basic morphological operations in IHMPs: $k = 10$.

Metrics	ARIMA	MMNN	MLP	MRL	D-IHMP	**E-IHMP**
MSE	2.1447e-5	9.7451e-5	9.6000e-3	1.0982e-4	3.8909e-5	**2.9857e-5**
MAPE	1.2400e-2	9.2000e-2	1.0103e-2	1.0214e-2	7.2277e-3	**6.2731e-3**
THEIL	1.4090	0.9498	0.9179	1.0397	0.6184	**0.5388**
ARV	0.1374	7.4749e-3	7.2875e-3	8.4926e-2	2.9930e-3	**2.2967e-3**
POCID	47.22	81.31	50.98	52.18	**100.00**	**100.00**
EF	18.4538	39.6756	26.2123	24.4409	61.4002	**64.6245**

Concluding Remarks

- We introduced the increasing hybrid morphological-linear perceptron (IHMP) with evolutionary learning.
- An automatic phase correction step is geared at eliminating time phase distortions.
- We conducted experiments using DJIA and S&P500.
- The IHMP outperformed competitive neural and statistical models in terms of 5 well-known performance measures and an evaluation function.
- The IHMP was able to cope with time phase distortions.
- The IHMP succeeds in modeling a combination of linear and nonlinear components by combining a linear module with a morphological or lattice-based module.
- Phase correction in IHMP adjusts the nonlinear component that enters the final prediction.