2-nd International Conference on Inductive Modelling – ICIM'2008

Some Fundamental Topics of Inductive Modeling

Yuriy V. Dzyadyk

International Center of Information Technologies and Systems, Academician Glushkov avenue, 40, Kyiv, 03680 Ukraine

iurius@inbox.ru iurius@i.com.ua

Kyiv, September 15–19, 2008

Agenda

1. Introduction
2. Problem of Unstability in Linear Modelling
3. Factor Analysis and Stabilization

   Method of Two Thresholds (MTT), or $(\beta,\ \gamma)$–Method
4. Stabilization Principle in Linear Modelling
5. Active Agent Models

   Cycles of Absorption and Reduction
6. Some Results

   Comparisons with Other Methods
7. Economic Criteria

The Most Interesting and Essential

3. Factor Analysis and Stabilization

   Method of Two Thresholds (MTT), or $(\beta,\ \gamma)$–Method

6. Some Results

   Comparisons with Other Methods

7. Economic Criteria

"Gentlemen! I must herald you the most disagreeable tidings ... " (N. V. Gogol')

It seems, the $(\beta,\ \gamma)$–method now is the best method of forecasting. I hope for discussion.

1. Introduction

2. Problem of Unstability in Linear Modelling

3. Factor Analysis and Stabilization

   Method of Two Thresholds (MTT), or $(\beta,\ \gamma)$–Method

4. Stabilization Principle in Linear Modelling

5. Active Agent Models

   Cycles of Absorption and Reduction

6. Some Results

   Comparisons with Other Methods

7. Economic Criteria

We have a series of observations

$$\mathbf{y} = (y_1,\ y_2,\ \ldots,\ y_n)^T. \tag{1}$$

We indend to find such set of independent variables $\{t_1,\ t_2,\ \ldots,\ t_l\}$ and their observations $\{\mathbf{t}_1,\ \mathbf{t}_2,\ \ldots,\ \mathbf{t}_l\}$ for building an inductive model

$$\hat{y} = f(t_1,\ t_2,\ \ldots,\ t_l). \tag{2}$$

We shall assume that model (2) is linear relative to some set $(x_1,\ x_2, \ldots, x_k)$ of **basic functions**, where each

$$x_i = \varphi_i(t_1,\ t_2,\ \ldots,\ t_l). \tag{3}$$

Usually we pick up these basic functions among $\binom{m+l}{m}$ monomials of all degrees $s \le m$ (as a rule, $m = 1$ or $m = 2$)

$$\prod_{j=1}^{s} t_{i_j},\ \ s = 0..m,\ \forall j : i_j \in \{1..l\}. \tag{4}$$

Some authors [1] mention as **basic** for **support functions** harmonic and logistic functions:

$$\sin(t_i),\ \cos(t_i),\ \frac{1}{1+e^{-t_i}}. \tag{5}$$

In other works as basic functions are using roots or fractional degrees, logarithms etc.

$$\sqrt[q]{t_i},\quad t_i^{p/q}\ (p \in \mathbb{Z},\ q \in \mathbb{N}),\quad \log(t_i)\ \text{ etc.} \tag{6}$$

---

1. Introduction

Let us remind that n denote the dimension of statistics. Thus from l vectors $\{\mathbf{t}_1,\ \mathbf{t}_2,\ \dots,\ \mathbf{t}_l\}$ we can compute k vectors $(\mathbf{x}_1,\ \mathbf{x}_2, \dots, \mathbf{x}_k)$ of a real n-dimensional space $\mathbb{R}^n$.

Now let $\mathbf{X}$ denote the matrix

(7) $$(\mathbf{x}_1 - \mu(\mathbf{x}_1),\ \mathbf{x}_2 - \mu(\mathbf{x}_2),\ \dots,\ \mathbf{x}_k - \mu(\mathbf{x}_k)) = \mathbf{X},$$

where

(8) $$\forall \mathbf{w} \in \mathbb{R}^n : \mu(\mathbf{w}) = \frac{1}{n}\sum_{t=1}^{n} w_i.$$

Note, that $\mathbf{x}_i$ are vector columns, μ_i are scalars. So, matrix $\mathbf{X}$ has n rows and k columns.

As model (2) $\hat{y} = f(t_1,\ t_2,\ \dots,\ t_l)$ is linear relative to the set $(x_1,\ x_2, \dots, x_k)$ of **basic functions**, there are such vector $\mathbf{a} = (a_1,\ a_2, \dots, a_k)$

(9) $$\hat{y} = \sum_{i=1}^{k} a_i x_i,$$

hence

(10) $$\hat{\mathbf{y}} = \sum_{i=1}^{k} a_i \mathbf{x}_i, \quad \mathbf{y} = \mathbf{X}\mathbf{a}.$$

1. Introduction

Let we search for a form $\mathbf{a}$ of the linear dependence $\mathbf{y} = \mathbf{X}\mathbf{a}$, where vector $\mathbf{y}$ and matrix $\mathbf{X}$ are known, vector $\mathbf{a}$ is sought. As well known, the heuristic (or symbolic) way

$$\mathbf{y} = \mathbf{X}\mathbf{a} \Longleftrightarrow \mathbf{X}^T\mathbf{y} = \mathbf{X}^T\mathbf{X}\mathbf{a} \Longleftrightarrow (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{y} = \mathbf{a} \tag{11}$$

leads to the same result that least-squares method (LSM):

$$\mathbf{a} = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{y}. \tag{12}$$

In order to investigate a pseudo-solution (12), let denote Gramian [3] matrix $\mathbf{X}^T\mathbf{X} = \mathbf{W}$. It is a square $(k,\ k)$ matrix.

Let be $\{\lambda_1, \lambda_2, \ldots, \lambda_k\}$ the set of all eigenvalues of $\mathbf{W}$ which are numbered so that $\lambda_1 \geq \lambda_2 \geq ... \geq \lambda_k$. Remind that for Gramian matrix $\mathbf{W}$ all λ_i are real nonnegative: $\lambda_i \geq 0\ \forall i$. Then:

$$\det \mathbf{W} = \lambda_1\lambda_2\ldots\lambda_k; \qquad \operatorname{trace} \mathbf{W} = \lambda_1 + \cdots + \lambda_k; \qquad \operatorname{cond} \mathbf{W} = \lambda_1(\lambda_k)^{-1}; \tag{13}$$

where $\operatorname{cond} \mathbf{W} = ||\mathbf{W}||\ ||\mathbf{W}^{-1}||$ denote a condition number, $||*||$ – any reasonable norm. Obviously, $\forall c : \operatorname{cond}(c\mathbf{W}) = \operatorname{cond} \mathbf{W}$, where c – an arbitrary constant.

Now, let us investigate the expression (12). It has no sense if

$$\det \mathbf{W} = 0 \Longleftrightarrow \lambda_k = 0 \Longleftrightarrow \operatorname{cond}\ \mathbf{W} = \infty.$$

Moreover, there is well-known fact: if cond $\mathbf{W}$ is near to infinity, i.e. λ_k is near to zero, then the solution $\mathbf{y} = \mathbf{X}\mathbf{a}$ is worthless for extrapolation or forecasting [2]. The obvious scholar example: the exact polynomial interpolation model, which is the solution of some matrix equation of the form $\mathbf{y} = \mathbf{X}\mathbf{a}$, has no sense beyond of interval of interpolation.

---

2. Problem of Unstability in Linear Modelling

Define the *measure of stability* of the matrix $\mathbf{X}$ as

$$\operatorname{stab}\mathbf{X} = \frac{k\lambda_k}{\operatorname{trace}(\mathbf{X}^T\mathbf{X})} = \frac{k\lambda_k}{\lambda_1 + \lambda_2 + \cdots + \lambda_k} \tag{14}$$

Obviously, for arbitrary matrix $\mathbf{X}$ and for any constant c:

$$0 \le \operatorname{stab}\mathbf{X} \le 1; \quad \forall c : \operatorname{stab}(c\mathbf{X}) = \operatorname{stab}(\mathbf{X}); \quad \operatorname{stab}\mathbf{X}\ \operatorname{cond}(\mathbf{X}^T\mathbf{X}) = \frac{k\lambda_1}{\lambda_1 + \lambda_2 + \cdots + \lambda_k} \tag{15}$$

and from unequalities $\lambda_1 \le \lambda_1 + \lambda_2 + \cdots + \lambda_k \le k\lambda_1$ we obtain:

$$1 \le \operatorname{stab}\mathbf{X}\ \operatorname{cond}(\mathbf{X}^T\mathbf{X}) \le k, \tag{16}$$

or

$$\operatorname{stab}\mathbf{X} \simeq [\operatorname{cond}(\mathbf{X}^T\mathbf{X})]^{-1} \tag{17}$$

The main problem is — *how to avoid inanity, insignificancy of the model* $\mathbf{y} = \mathbf{Xa}$ *beyond the neighbourhood of its construction domain*?

---

2. Problem of Unstability in Linear Modelling

1. Introduction

2. Problem of Unstability in Linear Modelling

3. Factor Analysis and Stabilization

Method of Two Thresholds (MTT), or $(\beta,\ \gamma)$–Method

4. Stabilization Principle in Linear Modelling

5. Active Agent Models

Cycles of Absorption and Reduction

6. Some Results

Comparisons with Other Methods

7. Economic Criteria

What is Factor Analysis?

Let reduce the Gramian matrix $\mathbf{X}^T\mathbf{X} = \mathbf{W}$ by orthogonal transformation $\mathbf{S}$ to such diagonal form $\mathbf{S}^T\mathbf{W}\mathbf{S} = \mathbf{D}$, that $d_{ii} = \lambda_i$ (let us remind that $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_k$). Then vectors (columns) of the matrix $\mathbf{X}\mathbf{S} = \mathbf{Z} = (\mathbf{z}_1,\ \mathbf{z}_2,\ \ldots,\ \mathbf{z}_k)$ are named as *factors* of $\mathbf{X}$.

Note, that $\forall i : \mu(\mathbf{z}_i) = 0$, so

(18) $$\forall(i, \mathbf{y}, c) : \langle \mathbf{y} + c, \mathbf{z}_i \rangle = \langle \mathbf{y}, \mathbf{z}_i \rangle.$$

By construction, first non-zero factors $(\mathbf{z}_1,\ \mathbf{z}_2,\ \ldots,\ \mathbf{z}_p)$ form an orthogonal basis of the linear enveloping space

(19) $$L = L(\mathbf{x}_1 - \mu_1,\ \mathbf{x}_2 - \mu_2,\ \ldots,\ \mathbf{x}_k - \mu_k),$$

where $p = \dim L \leq \min(k, n)$.

Thus, an arbitrary linear model $\hat{y}(x_1,\ x_2, \ldots, x_k)$ by transformation $\mathbf{S}$ may be represented in the form of

(20) $$\hat{y} = y_0 + y_1 z_1 + y_2 z_2 + \cdots + y_p z_p,$$

where

(21) $$\forall\ 0 < i \leq p,\ \ y_i = \frac{\langle \mathbf{y} - y_0,\ \mathbf{z}_i \rangle}{\langle \mathbf{z}_i,\ \mathbf{z}_i \rangle} = \frac{\langle \mathbf{y},\ \mathbf{z}_i \rangle}{|\mathbf{z}_i|^2} = \frac{\langle \mathbf{y},\ \mathbf{z}_i \rangle}{\lambda_i}.$$

---

3. Factor Analysis and Stabilization

Further, we define for every non-zero factor $\mathbf{z}_i$ two characteristics: *stability* $\mathrm{stab}\,(\mathbf{z}_i)$ and *essentiality* $\mathrm{essn}\,(\mathbf{y}, \mathbf{z}_i)$:

$$\mathrm{stab}\,(\mathbf{z}_i) = \frac{\lambda_i}{\mathrm{trace}\ (\mathbf{W})} = \frac{\lambda_i}{\mathrm{trace}\ (\mathbf{D})} = \frac{\lambda_i}{\lambda_1 + \lambda_2 + \cdots + \lambda_k} = \frac{\lambda_i}{\lambda_1 + \lambda_2 + \cdots + \lambda_p}; \tag{22}$$

of course,

$$\mathrm{stab}\,(\mathbf{Z}) = \mathrm{stab}\,(\mathbf{X}) = \min_i\ \mathrm{stab}\,(\mathbf{z}_i) = \mathrm{stab}\,(\mathbf{z}_k). \tag{23}$$

$$\mathrm{essn}\,(\mathbf{y}, \mathbf{z}_i) = \mathrm{corr}^2(\mathbf{y},\ \mathbf{z}_i) = \frac{\lambda_i y_i^2}{|\mathbf{y} - y_0|^2}, \tag{24}$$

where

$$\mathrm{corr}(\mathbf{y},\ \mathbf{z}_i) = \cos(\mathbf{y} - y_0,\ \mathbf{z}_i) = \frac{\langle \mathbf{y} - y_0,\ \mathbf{z}_i \rangle}{|\mathbf{y} - y_0||\mathbf{z}_i|} = \frac{\langle \mathbf{y},\ \mathbf{z}_i \rangle}{|\mathbf{z}_i|^2} \frac{|\mathbf{z}_i|}{|\mathbf{y} - y_0|} = \frac{y_i \sqrt{\lambda_i}}{|\mathbf{y} - y_0|}. \tag{25}$$

---

3. Factor Analysis and Stabilization

By $(\beta,\ \gamma)$–*reduction* of a model $\hat{y}$ by *method of two thresholds* (MTT), or $(\beta,\ \gamma)$–method, we call such diagonal projection $\mathbf{P}$ of the model (20), diag $\mathbf{P} = (s_1, s_2, \ldots, s_p)$

$$\hat{y}^s = y_0 + s_1 y_1 z_1 + s_2 y_2 z_2 + \cdots + s_p y_p z_p, \tag{26}$$

where $\forall\ i > 0,\ s_i = s_i(\beta,\ \gamma) = 0$, if $\operatorname{stab}(\mathbf{z}_i) < \beta$ or $\operatorname{essn}(\mathbf{y}, \mathbf{z}_i) < \gamma$, and $s_i = 1$ for all other factors.

Let's call factor $\mathbf{z}_i$ as *unstable* with the threshold β, or β–*unstable*, if $\operatorname{stab}(\mathbf{z}_i) < \beta$. Obviously, for every $\beta > 0$ there exists such $j \in \mathbb{N}$ that all β–*stable* factors form the set $(\mathbf{z}_1,\ \mathbf{z}_2,\ \ldots,\ \mathbf{z}_j)$, and all β–*unstable* factors form the set $(\mathbf{z}_{j+1},\ \ldots,\ \mathbf{z}_k)$.

Let's call factor $\mathbf{z}_i$ as *inessential* for variable $\mathbf{y}$ with the threshold γ, or γ–*inessential*, if $\operatorname{essn}(\mathbf{y}, \mathbf{z}_i) < \gamma$.

In these terms, $(\beta,\ \gamma)$–reduction is the obliteration of all β–*unstable* and γ–*inessential* factors.

We call $(\beta,\ 0)$–reduction as β–*stabilization*.

Let L be defined as (19), i.e. $L = L(\mathbf{x}_1 - \mu_1,\ \mathbf{x}_2 - \mu_2,\ \ldots,\ \mathbf{x}_k - \mu_k) = L(\mathbf{z}_1,\ \mathbf{z}_2,\ \ldots,\ \mathbf{z}_p)$, V is arbitrary subspace of L, and $\mathbf{P}$ is corresponding to V projection matrix, so that $\mathbf{XP}$ is the projection of $\mathbf{X}$ to V. We can substitute matrix $\mathbf{X}$ by matrix $\mathbf{XP}$ and apply all definitions: $\mathrm{cond}\,(\mathbf{P}^T\mathbf{WP})$, $\mathrm{stab}\,(\mathbf{XP})$ etc.

Let $\mathbf{XPS} = \mathbf{U} = (\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_w)$ be factors of $\mathbf{XP}$. We define *stability, essentiality and sufficiency* **of the subspace** V as

(27) $$\mathrm{stab}\,(V) = \min_{\mathbf{v}\in V}\ \mathrm{stab}\,(\mathbf{v}) = \min_{i\in\{1..w\}}\ \mathrm{stab}\,(\mathbf{u}_i) = \mathrm{stab}\,(\mathbf{u}_w);$$

(28) $$\mathrm{essn}\,(\mathbf{y},\ V) = \min_{\mathbf{v}\in V}\ \mathrm{essn}\,(\mathbf{y}, \mathbf{v}) = \min_{i\in\{1..w\}}\ \mathrm{essn}\,(\mathbf{y}, \mathbf{u}_i);$$

(29) $$\mathrm{suff}\,(\mathbf{y}, V) = \sum_{i=1}^{\dim V} \mathrm{essn}\,(\mathbf{y}, \mathbf{u}_i) = \sum_{i=1}^{\dim V} \mathrm{corr}^2(\mathbf{y},\ \mathbf{u}_i) = \frac{1}{|\mathbf{y} - y_0|^2} \sum_{i=1}^{\dim V} \left[\frac{\langle \mathbf{y}, \mathbf{u}_i\rangle}{|\mathbf{u}_i|}\right]^2 .$$

Definition. We call subspace V as:
β–*stable*, if $\mathrm{stab}\,(V) > \beta$;
γ–*essential*, if $\mathrm{essn}\,(V) > \gamma$;
δ–*sufficient*, if $\mathrm{suff}\,(\mathbf{y}, V) > 1 - \delta$.
If β, γ or δ are implicit as fixed, we call subspace V simply as *stable*, *essential* or *sufficient*.

As a rule, implicit values are: $\beta = 10^{-3}$, $\gamma = 0$ or $\gamma = 10^{-4}$, $\delta = 5\%$ or $\delta = 10^{-2}$.

---

4. Stabilization Principle in Linear Modelling

Let $\sigma = \{i_1,\ i_2,\ \dots,\ i_\sigma\}$ be any subset of $K = \{1,\ 2,\ \dots,\ k\}$. Denote by $V(\sigma,\ X)$, $V(\sigma,\ Z)$ the linear enveloping space of vector columns $\mathbf{x}_i\mu(\mathbf{x}_i)$, $\mathbf{z}_i$, respectively, for all $i \in \sigma$.

Hypothesis. The *essence* of GMDH is the projection of L on such stable and essential subspace $V(\sigma,\ X)$, which is the most *sufficient*.

Definition. By general $(\beta,\ \gamma)$–method, or stabilization principle, we call a search of the solution of the next optimization problem: to find in some set $S(L) \subseteq 2^L$ of subspaces of L such β–stable and γ–essential subspace V which is the most sufficient:

$$\{\text{stab}\,(V) > \beta\}\ \&\ \{\text{essn}\,(\mathbf{y},\ V) > \gamma\}\ \&\ \{\text{suff}(\mathbf{y},\ V) = \max\}\ \&\ \{V \in S(L) \subseteq 2^L\}, \tag{30}$$

where 2^L is the set of all subspaces of L.

Examples:
(a) $S(L) = \{V(\sigma,\ X) : \sigma \in 2^K\}$;
(b) $S(L) = \{V(\sigma,\ Z) : \sigma \in 2^K\}$;
(c) $S(L) = 2^L$.

4. Stabilization Principle in Linear Modelling

1. Introduction
2. Problem of Unstability in Linear Modelling
3. Factor Analysis and Stabilization

   Method of Two Thresholds (MTT), or $(\beta,\ \gamma)$–Method
4. Stabilization Principle in Linear Modelling

5. Active Agent Models

   Cycles of Absorption and Reduction

6. Some Results

   Comparisons with Other Methods
7. Economic Criteria

As a rule, neither $(\beta,\ \gamma)$–method of two thresholds (MTT) nor general $(\beta,\ \gamma)$–method gives us the δ–sufficient solution $\mathbf{a}$. So, model $\mathbf{y} = \mathbf{Xa}$ would be stable but not sufficiently exact.

The next stage is *absorption* of other variables (or, in economis terms, activities) in the information space (as a rule, in data bases in Internet). This part of inductive modelling is very interesting and complicated. It deals with terms of constructing intelligent agents (CIA) [4], artificial intelligence (AI), multiagent systems (MAS) et cætera.

But most of these topics are not yet included in the Inductive Modelling Theory.

Let give some initial definitions fit for the field of inductive modelling.

Active model is the model which are searching all relevant activities $(t_1,\ t_2,\ \ldots,\ t_l)$ are necessary for its functioning (and statistics of these activities) in the information space (mainly in Internet).

Active agent model is the model which are searching in Internet by the instrumentality of intelligent search agents. Synonym: active intelligent model.

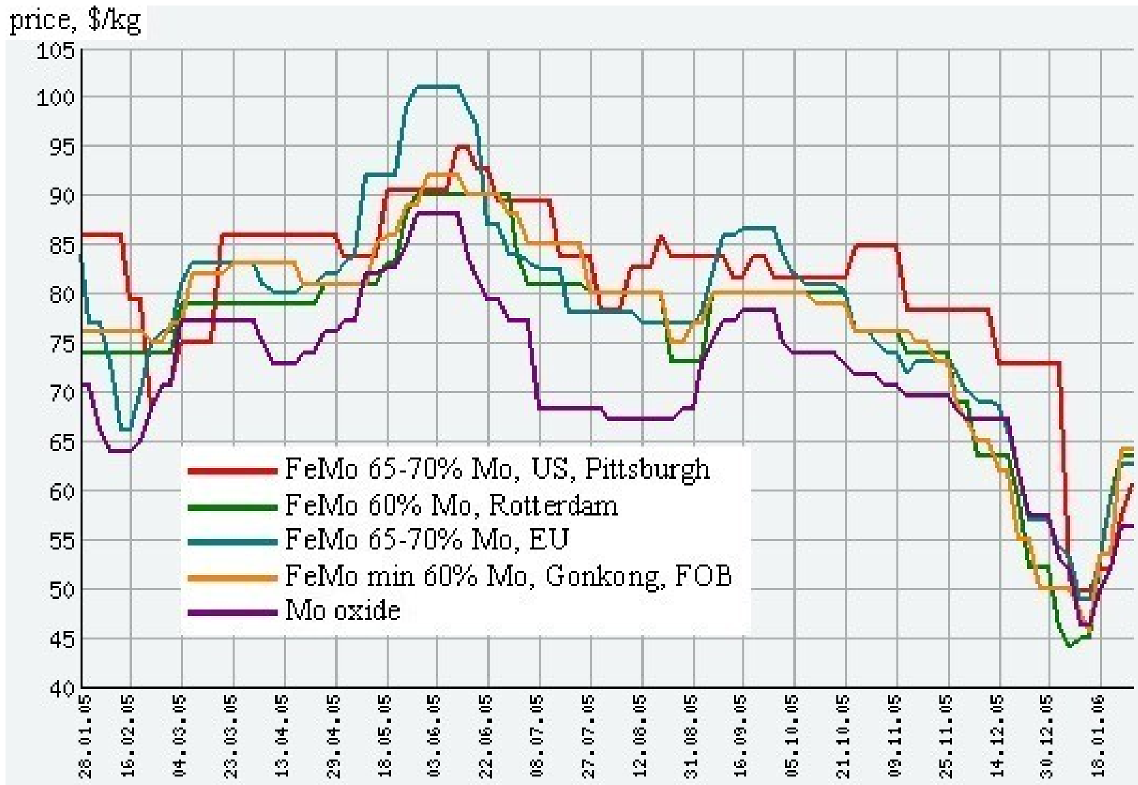
Active forecasting is the forecasting by the instrumentality of active intelligent model.

It may be very interesting to investigate the image of every cycle of absorption and reduction on the two dimensional plane "stability – sufficiency". At the first look, this image is some similar to Carnot cycle, but this impression is superficial.

---

5. Active Agent Models. Cycles of Absorption and Reduction

Firstly, the method of two thresholds was successfully used for modelling and forecasting of ferromolybdenum and molybdenum prices. Especially interesting was forecast of extremely quick jumping monthly prices in 2004–07. Even the possibility of any forecast under these data was under question [5].



---

6. Some Results

For monthly moving forecasting of molybdenum prices in 2005-07, in the period 2004–05 years in Internet were selected $l = 9$ activities $\{t_1,\ t_2,\ \ldots,\ t_l\}$:
t_1, t_2 – molybdenum export and import prices, which were calculated from data given in tables on Internet site [6];
t_3 – import prices on unwrought copper, were calculated in the same way from [6];
$t_4,\ t_5,\ \ldots,\ t_9$ – world stainless steel prices of 6 different grades [7] (see Tab.).

Note, that these 9 activities were selected only after 4-th cycle of absorption and reduction. *All* previous activities went to the wastepaper basket (or to recycle bin).

From $l = 9$ activities $\{t_1,\ t_2,\ \ldots,\ t_l\}$ we form a forecasting variable $y(\tau)$ and $k = 12$ input basic variables $\{x_1,\ x_2,\ \ldots,\ x_k\}$:

$y(\tau) = t_1(\tau + 1);$

$x_1(\tau) = t_1(\tau - 2),\ x_2(\tau) = t_1(\tau - 1),\ x_3(\tau) = t_1(\tau),$
$x_4(\tau) = t_2(\tau - 1),\ x_5(\tau) = t_2(\tau),$
$x_i(\tau) = t_{i-3}(\tau),\ i = 6..12.$

---

6. Some Results

Tab. 1: MEPS – World Stainless Steel Product Prices ($ US/tonne).

Date	Hot Rolled Plate		Cold Rolled Coil		Drawn Bar	
	Grade 304	Grade 316	Grade 304	Grade 316	Grade 304	Grade 316
Jan 04	2117	2915	2137	2919	2376	3111
Feb 04	2367	3206	2372	3222	2556	3349
Mar 04	2468	3389	2484	3399	2639	3480
Apr 04	2540	3451	2509	3409	2697	3546
...	...	...	...	...	...	...
Sep 05	2768	4933	2487	4714	2968	5225
Oct 05	2714	4746	2434	4534	2895	4973
Nov 05	2665	4750	2389	4531	2803	4925
Dec 05	2506	4578	2217	4341	2607	4705

6. Some Results

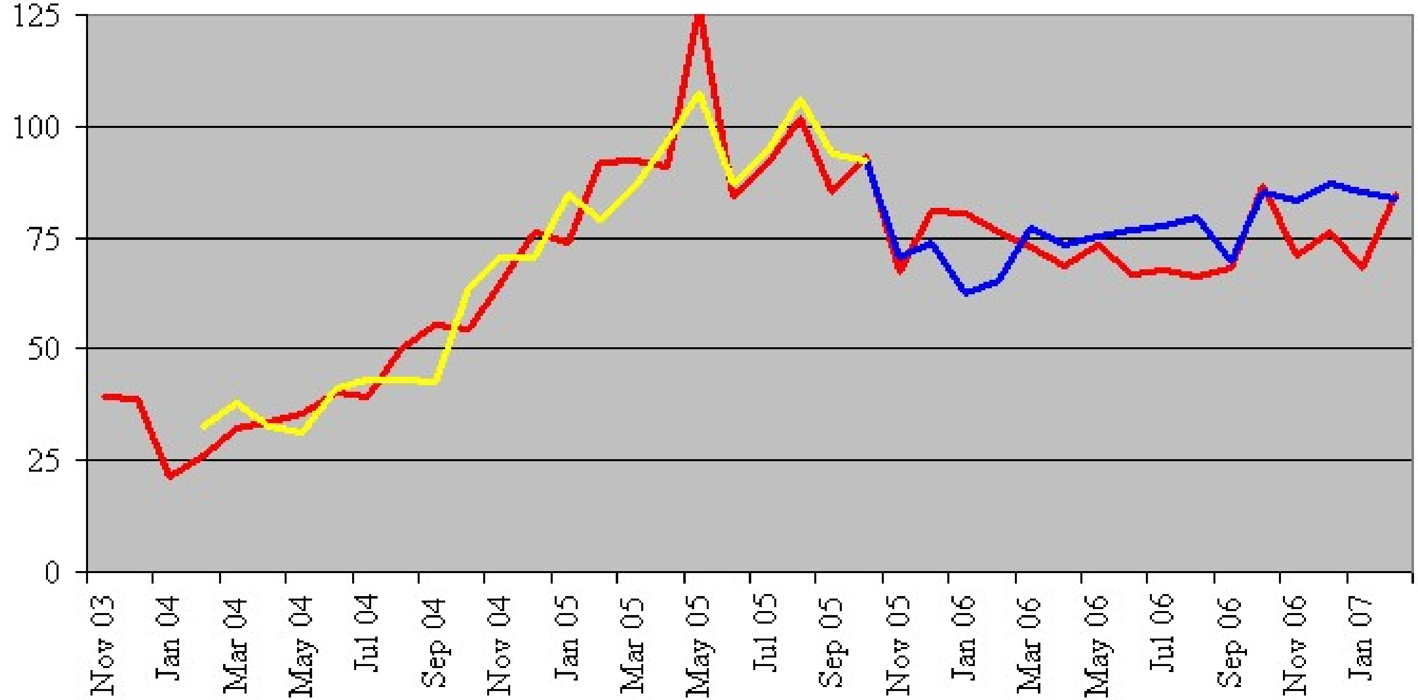


Figure 1: Graphics of the actual values of molybdenum prices (red line), simulated (yellow line) and forecasted values (blue line)

---

6. Some Results

Tab. 2: Results of forecasting of Molybdenum price ($ US/kg)

Month	Price		Diffe-	Change	
	Fact	Forecast	rence	Fact	Forecast
Oct 05	93.51				
Nov 05	67.45	70.36	2.91	**-26.05**	**-23.15**
Dec 05	81.06	73.88	-7.18	13.61	6.43
Jan 06	80.58	62.40	-18.18	-0.48	-18.66
Feb 06	76.19	65.50	-10.69	-4.39	-15.08
Mar 06	73.15	77.08	3.94	-3.04	0.90
Apr 06	68.49	73.37	4.88	-4.66	0.22
May 06	73.38	75.50	2.12	4.89	7.01
Jun 06	66.83	76.55	9.72	-6.55	3.17
Jul 06	67.55	77.46	9.91	0.72	10.63
Aug 06	66.29	79.73	13.45	-1.26	12.18
Sep 06	68.37	69.60	1.23	2.08	3.32
Oct 06	86.76	85.23	-1.53	**18.39**	**16.86**
Nov 06	71.01	83.53	12.52	-15.75	-3.23
Dec 06	76.35	87.05	10.70	5.33	16.03
Jan 07	68.18	85.02	16.83	-8.16	8.67
Feb 07	84.67	84.02	-0.65	**16.48**	**15.84**

Other example is based on medical data from Motol hospital in Prague. It was proposed me by P. Kordik to compare $(\beta,\ \gamma)$–method with results of [8] which were obtained by means of Group of Adaptive Models Evolution (GAME).

Data of the Tab. 3 and Fig. 2 demonstrate the advantage of MTT over GMDH and GAME. It is worthy of note that GAME needs huge training data (some hundreds of observations) [8], when MTT needs about 10-20 observations.

Tab. 3: Comparison of the $(\beta,\ \gamma)$–method (MTT) with GMDH and GAME [8] on examples of molybdenum price forecasting (row Mo) in 2005–07 and forecast of CO2 in brain [8] (row CO2): RMS error of prediction

example\method	MTT	GMDH, FPE	GMDH, AR	GMDH, AC	GAME
CO_2	0,0380	–	–	0,0704	0,0386
Mo	9,62	20,22	12,51	–	–

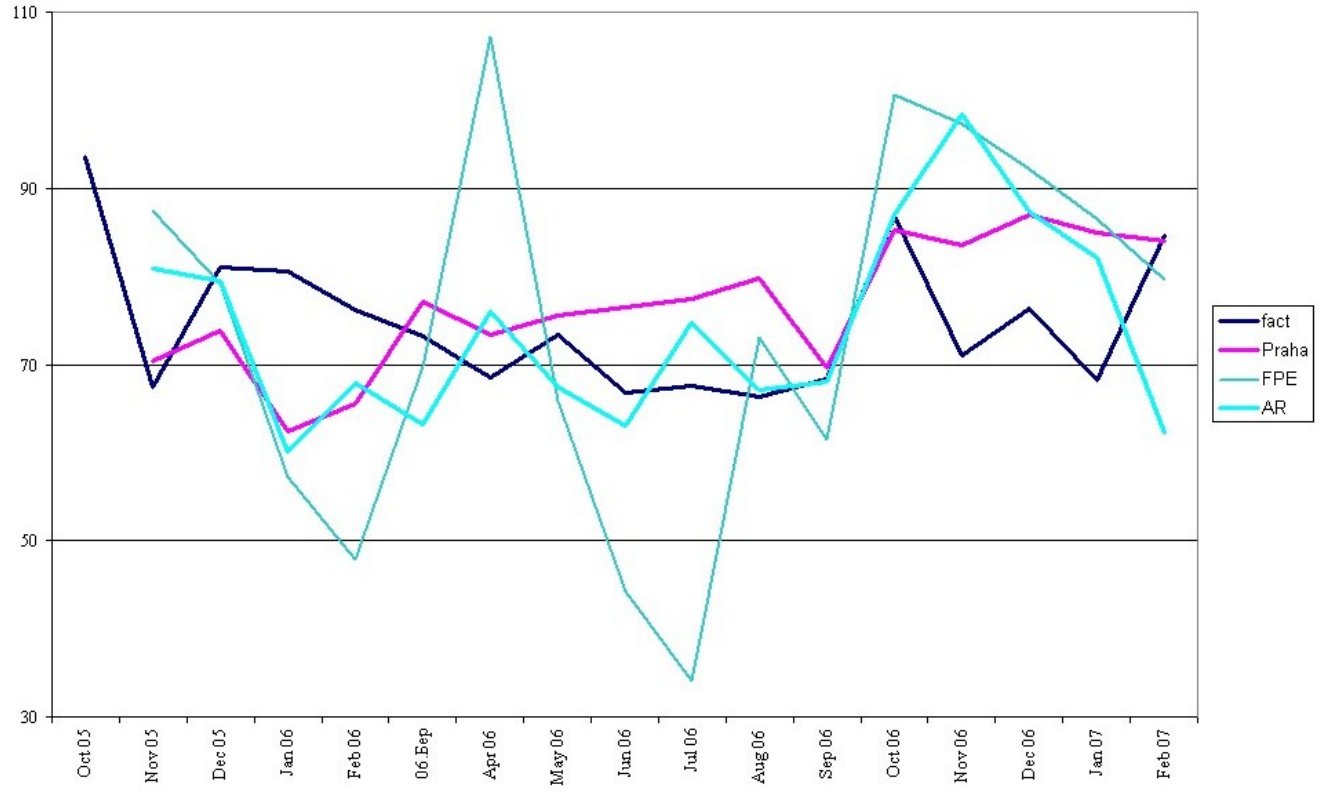


Fig. 2: Graphics of the actual values of molybdenum prices (line fact), and forecast values, computered by $(\beta,\ \gamma)$–method (line Praha) and GMDH (lines FPE and AR)

---

6. Some Results. Comparisons with Other Methods

7. Economic Criteria

Let y be a price for some good. We can buy this good month by month, then expenses equal to

$$Z_0 = \sum_{i=1}^{n} y_i. \tag{31}$$

On the end of the period we can see the least price ymin. If we buy all good by the least price, expences would equal to

$$Z_{\min} = n y_{\min}. \tag{32}$$

Of course, we are not omniscient. But if we have a forecast by method M, we can use it in some way W for optimal planning of purchase. Then we can compute corresponding expenses $Z_{M,\ W}$:

$$Z_{M,\ W} = \sum_{i=1}^{n} p_i(M,\ W) y_i, \tag{33}$$

where $p_i(M,\ W)$ is the advice for buying, formed using forecast method M and way of using W. It gives some types of economic criteria. E.g., relative criterion

$$\frac{Z_0 - Z_{M,\ W}}{Z_0 - Z_{\min}}. \tag{34}$$

for molybdenum price forecast by MTT equals about 68%.

Let some plant consumes 20 tons of molybdenum per month and use forecast (Table 2) plus some simple heuristic rule for decision making of its purchase. Then absolute economy of this simulated plant would equal about $ 1,9 mln over period Nov 2005 – Feb 2007 (16 monthes).

7. Economic Criteria

References

[1] Ivakhnenko A. G., Ivakhnenko G. A.: The Review of Problems Solvable by Algorithms of the Group Method of Data Handling (pdf)

[2] Koppa Yu. V., Stepashko V. S.: A Comparison of the Forecasting Properties of Regression Types Models versus GMDH (pdf) (*in Russian*)

[3] Gramian matrix.
Jørgen Pedersen Gram.

[4] Joseph P. Bigus, Jennifer Bigus. Constructing Intelligent Agents Using Java: Professional Developers Guide. John Wiley & Sons, 2001, second edition. ISBN 0471191353

[5] Private letter of Pavel Kordik.

[6] Metals Statistics, U.S. Metals Trade

[7] MEPS (International) Ltd. – Independent Steel Industry Analysts, Consultants, Steel Prices, Reports and Publications. World Stainless Steel Product Prices

[8] Josef Bouška, Pavel Kordik. Time Series Prediction by means of GMDH Analogues Complexing and GAME

Bibliography