Computer Science Journal of Moldova, vol.**5**, no.2(14), 1997

Construction of complicated multifactor models of optimal form on the basis of statistical data, when the physical essence of phenomenon is known

E. Livovsky G. Cioban

Abstract

The work contains computer-aided method of contruction of multifactor and complicated models on the basis of experimental data. The models are non-linear but traditional (constructed with the help of the method with the preliminary transformation of the matrix of the initial data and the subsequent reverse transformation of the results). The models can have different form linear, multiplicative, exponent and so on. After the construction the models are improved using method of Givis and Hook, which improves their statistical characteristics and the guality of representation of the results of experiment. This work is in a certain way methodological. We have shown in it that having enough experimental data we can improve the quality of the models and formulas even those which already exist in school textbooks. On the other hand it is shown here that when knowing the essence of the phenomenon it is possible to choose an optimal model from the multitude of regressional models.

It's determinate, that while constructing models of multifactor regression, the number of these models may be infinite.

However if we analyse the physical essense of the phenomenon thoroughly, an optimal model can be built. To demonstrate the posibilities of the method we have specially chosen the models that have been studied by specialists about a hundred years. The form of these models is established but it is interesting to check on the basis of the experimental material how close the values of the empirical coefficients in use are to the true ones. Two models of mechnical characteristics of concrete have been chosen. The first model, the formula for the modulus deformations of concrete is simpler and contains only one exponent. The second one which is much more complicated presens itself as a model of creep measure and contains two exponents.

The Chair of Building Constructions of Technical University of Moldova has been for 20 years doing work on construction of mathematical models on the basis of experimental data [1]. We have accumulated great packet of software which allows to carry out all necessary work on processing of the results of experiments.

This work includes preliminary processing of experiments: sifting of wild experiments checking of the hypothesis of normality of distributions graphic analysis of the results of experiments, that allows to make prognosis about the form of mathematical models.

A special program that we have PRE.FOR allows to transform initial data so, as to be able to obtain mathematical models of any form and complication.

We would like to note (point out) that the programs and computers we have,allow to construct multyfactor mathematical models with any number of factors and with any dimension of the matrix of initial data.

After the construction the models are thouroughly analysed and then we try to improve the quality of predication of the models, e.g. using Hook-Givis method. is given below.

Part one.
Construction of the simple model with one exponent

Dozens of models for calculation of the deformation modulus of concrete E_b have been proposed in world practice of calculation of mecanical proprieties of concrete. We hope that one of the most universal models is the model used in Maslov-Arutiunian theory of creep of the concrete.

This model is:

$$E_b = Eo\left[1 - \beta e^{-\alpha_1 \tau}\right] \tag{1}$$

where E_o is the limiting value of the deformation modulus, β, α_1 are empirical coefficients.

It is evident that E_o depends on many factors (more than 30 ones of them are known). We have a lot of experimental rezults obtained by 22 specialists in the field [2] E_o is found to be determined by 10 factors.

These are:

1. Mass of agregate Ag in $1m^3$ of concrete in $\frac{kg}{100}$.

2. Mass of cement C in $1m^3$ of concrete in kg.

3. Humidity H of the medium in %.

4. Scale factor S in m^{-1}.

5. Water-cement relation W/C in %.

6. Normal consistence CN of the cement in %.

7. Deformation modulus of the stone of agregate E in $MPa/100$.

8. Cement rezistance R_c in Mpa.

9. Maximal size of the agregate d_{max} in mm.

Besides, values of τ are given to determine E_b in $MPa/100$ as a function of the age of concrete τ (in days).

Process of modelling was performed by steps: at the beginning linear models were constructed and examined, then the multiplicative ones and finally the more complicated (multiplicative-exponential) models were examined.

The first step. Formation of the linear model with 10 factors based on 265 experiments

Results for calculation

The linear regression has 11 columns (function Y + 10 factors X_j and 265 lines.

Y	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	X_{10}
E_b	Ag	C	W	S	A/C	τ	CN	E	Rc	d_{max}

Means:

285.88; 18.68; 344.94; 78.95; 41.34; 54.68; 48.55; 25.29; 459.02; 31.40; 21.04

Correlation coef.:

1	+1.00	+0.08	−0.05	−0.11	+0.19	−0.42	+0.33	+0.18	−0.11	+0.31	−0.03	
2	+0.08	+1.00	−0.71	−0.28	−0.08	+0.37	−0.01	+0.01	+0.20	−0.34	+0.29	
3	−0.05	−0.71	+1.00	+0.33	−0.05	−0.59	−0.02	+0.11	+0.05	+0.17	−0.21	
4	−0.11	−0.28	+0.32	+1.00	−0.28	−0.01	−0.05	−0.17	+0.05	+0.06	+0.19	
5	+0.19	−0.08	−0.05	−0.28	+1.00	−0.22	+0.16	+0.11	−0.11	+0.01	−0.36	
6	−0.42	+0.37	−0.59	−0.01	−0.22	+1.00	−0.08	−0.35	+0.29	−0.25	+0.10	(2)
7	+0.33	−0.01	−0.02	−0.05	+0.16	−0.08	+1.00	−0.01	+0.03	+0.12	−0.09	
8	+0.18	+0.01	+0.11	−0.17	+0.11	−0.35	−0.01	+1.00	−0.18	+0.24	−0.08	
9	−0.11	+0.20	+0.05	+0.05	−0.11	+0.30	+0.02	−0.18	+1.00	−0.30	+0.12	
10	+0.32	−0.34	+0.17	+0.06	+0.01	−0.25	+0.12	+0.24	−0.30	+1.00	−0.09	
11	−0.03	+0.29	−0.21	+0.19	−0.36	+0.10	−0.09	−0.08	+0.12	−0.09	+1.00	

Vector RY:

+0.084; -0.051; -0.106; +0.192; -0.420; +0.333; +0.180; -0.114; +0.316; -0.034.

Determinant of matrix $RX = 0.0066$

Standard regression coef.:

+0.125; -0.469; +0.069; -0,002; -0.707; +0.226; -0.049; +0.170; +0.291; -0.093.

Natural regression coef.:

+6.661; -0.517; +0.296; -0.015; -5.490; +0.249; -2.289; +0.122; +2.974; -0.868.

Particular correlation coef.:

+0.102; -0.300; +0.082; -0.003; -0.53; +0.292; -0.061; +0.189; +0,331; -0.108.

t-values: $t_{Ag} = +1.603$; $t_C = -4.924$; $t_H = +1.289$; $t_S = -0.039$; $t_{W/C} = -9.685$; $t_\tau = +4.774$; $t_{CN} = -0.953$; $t_E = +3.008$; $t_{d_{R_C}} = +5.500$; $t_{d_{max}} = -1.705$.

Absolute term of equation +531.921.

Multiplication correlation coef. $R = 0.682$.

Sum of squares of rezidnes $SS_r = 1116653.000$.

F-value $F = 1.872$.

The general aspect of linear model:

$$\begin{aligned} E_b &= 531.92 + 6.66Ag - 0.52C + 0.30H - 0.02S - 5.49W/C \\ &\quad +0.25\tau - 2.29CN + 0.12E + 2.97R_c - 0.87d_{max} \end{aligned} \qquad (3)$$

Only 5 factors $(C, W/C, \tau, E, R_c)$ have statistical significance (criterium $t > 2$).

The second step.
A linear model with 8 factors

This model which had not much difference from the model (3). In this model two factors were eliminated: S and CN.

The third step

A multiplicative model was formed with the same 8 factors.

$$\begin{aligned} E_b &= 28928.13 \cdot Ag^{0.26439} \cdot C^{-0.72647} \cdot H^{0.01988} \cdot (W/C)^{-1.16740} \\ &\quad \cdot \tau^{0.16759} \cdot E^{0.23252} \cdot R_c^{0.37634} \cdot d_{max}^{0.02634} \end{aligned} \qquad (4)$$

Statistical values.

The sum of squares of rezidnes $SS_r = 1201161.0$; coefficient of multiplicative correlation $R = 0.73$; Fisher's criterium $F = 2.15$.

In this model only the factors Ag, H and d_{max} are statistically unvaluale $(t < 2)$.

The fourth step

29 experiments were eliminated and the multiplicative model with the same 8 factors was formed:

$$\begin{aligned} E_b &= 11781.18 \cdot Ag^{0.61250} \cdot C^{-0.73418} \cdot H^{0.02926} \cdot (W/C)^{-1.20270} \\ &\cdot \tau^{0.15894} \cdot E^{0.21205} \cdot R_c^{0.43375} \cdot d_{max}^{0.02007} \end{aligned} \quad (5)$$

Here two factors H and d_{max} are statistically unvaluable.

Statistical values.

The sum of squares of residnes $SS_r = 660418.9$; coefficient of multiplicative correlation $R = 0.80$; Fisher's criterium $F = 2.76$.

The fifth step

A multiplicative model with only 6 factors was formed

$$(Ag, C, W/C, \tau, E, R_C)$$

— all of them being statistically valuable:

$$\begin{aligned} E_b &= 8989.86 \cdot Ag^{+0.66890} \cdot C^{-0.70777} \cdot (W/C)^{-1.18207} \\ &\tau^{+0.15612} \cdot E^{+0.21515} \cdot R_c^{+0.44704} \end{aligned} \quad (6)$$

Statistical values.

The sum of squares of rezidnes $SS_r = 675237.2$; coefficient of multiplicative correlation $R = 0.80$; Fisher's criterium $F = 2.75$.

The sixth step

A multiplicative-exponential model with the same 6 factors is formed:

$$E_b = E_o\left(1 - e^{-0.3\times 1.01752\tau}\right) \quad (7)$$

where

$$E_o = \prod_{j=0}^{p} A_j^{x_j} \tag{8}$$

where p is number of factors X

$$\begin{aligned} E_o &= 7815.33 \cdot Ag^{+0.86207} \cdot C^{-0.69998} \cdot (W/C)^{-1.33895} \\ &\quad \cdot E^{+0.30234} \cdot R_c^{+0.50441} \end{aligned} \tag{9}$$

Statistical values.

t-values: $t_{Ag} = +3.921$; $t_C = -6.838$; $t_{A/C} = -15.409$; $t_E = +4.176$; $t_{R_c} = +10.954$; $t_\tau = +14.655$.

The sum of squares of rezidnes $SS_r = 390596.2$; coefficient of multiplicative corelation $R = 0.86520$; Fisher's criterium $F = 3.977$.

It's the best model for calculation of modulus of deformation of concrete.

One could expect that the model obtained by statistical methods could not be improved. However, it must be mentioned that the exponential coefficients are selected by trial method and can be specified. Hence, the value $t_\tau = 14.655$, $(t >> 2)$ can be also used.

The last, seventh step

The Givis-Hook method was used to optimize the model. Afther 4450 iteration the sum of squares of residnes was desceased to $SS_r = 358837.5$. So, the sum of squares was decreased 3 times.

$$\Delta = \frac{1116653.0}{358837.5} = 3.11. \tag{10}$$

The multiplicative-exponential model for calculation of deformation modulus E_b is:

$$E_b = E_o\Big(1 - e^{-0.330953}\Big). \tag{11}$$

where

$$E_o = 7824.288 \cdot Ag^{+0.65884} \cdot C^{-0.5870556} \cdot (W/C)^{-1.129662} \cdot E^{+0.2033501} \cdot R_c^{+0.4217347} \quad (12)$$

The sum of squares of residnes $SS_r = 358837.5$.

Part two. Construction of the complicated model with two exponente. Model of creep of concrete

Model of measure of creep of concrete can be shown as a product of multiplication of functions:

$$C_{(t,\tau)} = \Theta_\tau \times \Theta_{(t-\tau)} \quad (13)$$

where

$$\Theta_{tau} = C_o e^{-\gamma\tau} \quad (14)$$

$$\Theta_{(t-\tau)} = 1 - e^{-\gamma_1(t-\tau)} \quad (15)$$

$$C_o = \prod_{j=0}^{p} A_j^{x_j} \quad (16)$$

where p is number of factors X

Data for analyses. Function and factors

y is the function, specific relative deformations of creep of concrete, $C_{(t,\tau)}$, $((kg/cm^2)^{-1})$.

$x_1 - Ag-$ mass of aggregate of concrete, $kg/100$,
$x_2 - C-$ mass of cement, kg,
$x_3 - H-$ humidity of air, %,
$x_4 - S-$ factor of scale, cm,
$x_5 - W/C-$ water-cement relation, %,
$x_6 - \tau-$ age of concrete, day,
$x_7 - (t-\tau)$ time of change of the deformation, day.

The first problem. Simple linear regression

Initial (experimental data - function and factors) - matrix (8 columns, 260 lines).

Results for calculation

Means:

77.079; 18.278; 347.431; 70.312; 14.880; 54.180; 37.950; 762.462.
Correlation coef.:

1	+1.000	+0.071	−0.297	−0.340	−0.104	+0.613	−0.146	+0.366
2	+0.071	+1.000	+0.008	−0.073	−0.070	+0.097	+0.024	+0.334
3	−0.297	−0.008	+1.000	+0.172	−0.056	−0.592	−0.059	−0.338
4	−0.340	−0.073	+0.172	+1.000	−0.081	+0.044	−0.049	−0.246
5	−0.104	−0.070	−0.056	−0.081	+1.000	−0.079	−0.022	−0.016
6	+0.613	+0.097	−0.592	+0.044	−0.079	+1.000	−0.051	+0.200
7	−0.146	+0.024	−0.059	−0.049	−0.022	+0.051	+1.000	+0.077
8	+0.366	+0.334	−0.338	−0.246	−0.016	+0.200	+0.077	+1.000

(17)

Vector RY
+0.071; -0.297; -0.340; -0.104; +0.613; -0.146; +0.366.
Determinant of matrix $RX = 0.426$.
Standard regression coef.
-0.127; +0.307; -0.383; -0.066; +0.774; -0.206; +0.279.
Natural regression coef.
-2.796; +0.239; -1.203; -0.612; +3.111; -0.239; -0.018.
Particular correlation coef.
-0.196; +0.361; -0.524; -0.109; +0.715; -0.329; +0.378.
t-values: $t_{Ag} = -3.129$; $t_C = +6.061$; $t_H = -9.631$; $t_S = -1.719$; $tW/C = +16.00$; $t_\tau = -5.460$; $t_{t-\tau} = +6.393$.
Absolute term of equation -34.110
Multiplication correlation coef. $R = 0.802$.
Sum of squares of residnes $SS_r = 287013.300$.
F-value $F = 2.808$.

The general aspect of linear model

$$C_{(t,\tau)} = -34.110 - 2.796Ag + 0.239C - 1.203H - 0.612S$$

$$+3.111W/C - 0.239\tau + 0.018(t-\tau) \tag{18}$$

The second problem: multiplication model

$$\begin{aligned} C_{(t,\tau)} &= +0.01484 \cdot Ag^{-1.28943} \cdot C^{+1.07747} \cdot H^{-1.08849} \\ &\cdot S^{-0.15218} \cdot W/C^{+2.36272} \cdot \tau^{-0.19382} \cdot (t-\tau)^{+0.32849} \end{aligned} \tag{19}$$

Statistical values

t-values: $t_{Ag} = -10.661$; $t_C = +9.536$; $t_H = -14.513$; $t_S = -2.146$; $t_{W/C} = +25.516$; $t_\tau = -12.146$; $t_{(t-\tau)} = 17.968$.

$SS_r = 160252.200$; $R = 0.92427$; $F = 6.86239$.

The third problem: multiplication-exponential model

$$\begin{aligned} C_{(t,\tau)} &= 0.76583 \cdot Ag^{-1.15646} \cdot C^{+0.80148} \cdot H^{-1.21177} \cdot S^{-0.19411} \\ &\cdot W/C^{+2.30500} \cdot e^{-0.0063972\tau} \cdot \left[1 - e^{-0.00400075(t-\tau)}\right] \end{aligned} \tag{20}$$

Statistical values

t-values: $t_{Ag} = -10.753$; $t_C = +7.874$; $t_H = -17.846$; $t_S = -3.008$; $t_{W/C} = +27.549$; $t_\tau = -14.875$; $t_{(t-\tau)} = 19.913$.

$SS_r = 125089.800$; $R = 0.93735$; $F = 8.23870$.

The fourth problem: improvement of model

$$\begin{aligned} C_{t,\tau} &= 0.28714 \cdot Ag^{-1.28145} \cdot C^{+0.97458} \cdot H^{-1.19028} \cdot S^{-0.05284} \\ &\cdot W/C^{2.29649} \cdot e^{-0.00567\times\tau} \cdot \left[1 - e^{-0.00494\times(t-\tau)}\right] \end{aligned} \tag{21}$$

or

$$C_{(t,\tau)} = \Theta_\tau \times \Theta_{(t-\tau)} \tag{22}$$

where

$$\Theta_{tau} = C_o e^{-0.00567\times\tau} \tag{23}$$

$$\Theta_{(t-\tau)} = 1 - e^{-0.00494\times(t-\tau)} \tag{24}$$

$$C_o = \prod_{j=0}^{p} A_j^{x_j} = 0.28714 \cdot Ag^{-1.28145} \cdot C^{+0.97458} \cdot H^{-1.19028} \cdot S^{-0.05284} \cdot W/C^{2.29649} \qquad (25)$$

The values γ and γ_1 that we have obtained are very different from the already known ($\gamma = 0.012, \gamma_1 = 0.006$), therefore we have managed to get a better model for calculation $C_{(t,\tau)}$.

Conclusions

The following is shown in the article:

1. How an optimum regression model can be built, if the phisical meaning of the studied phenomenon is known.
2. How, having enough experimental data, we can ameliorate the well-known empirical formulae even those included in school manuals many years ago.

References

[1] E.Livovsky, Statistical Methods for Construction of Empirical Formulas, Moscow, “High Shcool”, 1988, 240 pp.

[2] E.Livovsky, Passive and Active Experiments for Study of mecanicals Proprieties of Concrete, Kishinau, “Cartea Moldoveneasca”, 1970, 180 pp.

Eugen Livovsky, Galina Cioban,
Technical University of Moldova,
Kishinev, Moldova

Received April 1, 1997