

TVORBA REGRESNÝCH MODELOV S VYUŽITÍM PROGRAMOVÉHO SYSTÉMU MATLAB

Gombár Miroslav, Hloch Sergej, Valíček Ján, Pella Alexander

Resumé

Príspevok sa zaoberá teoretickým popisom tvorby matematicko – štatistických modelov s využitím metódy najmenších štvorcov. Tvorba regresných modelov má pri skúmaní závislostí rôznych technologických a technických systémov a procesov mimoriadny význam, jednak z hľadiska pochopenia ich podstaty, možnosťou predikcie jeho správania a jednak z hľadiska jeho možnej optimalizácie.

Abstract

The paper deals with theoretical description of mathematical-statistic models creation by means of least square method. Regression is used to model relationships between variables and determine the magnitude of those relationships. Regression model creation has remarkable importance, for understanding of their essence, with possibility of the prediction of the behavior and from its possible optimization.

1 Štatistická analýza závislostí

Pri sledovaní súvislostí vecí a javov sa stretávame s rozličnými formami závislostí. V podstate rozoznávame súvislosti nepodstatné a na druhej strane súvislosti podstatné. Za nepodstatné súvislosti považujeme také, pri ktorých vonkajšia súvislosť (t.j. skutočnosť, že určité javy spoločne vznikajú, majú rovnaký priebeh, alebo po sebe nasledujú v pravidelne sa opakujúcej postupnosti) nemá vnútornú nevyhnutnosť. Podstatné súvislosti sú také, pri ktorých vonkajšia súvislosť je výrazom určitej vnútornej nevyhnutnosti.

Charakteristickou formou podstatnej súvislosti je príčinná alebo kauzálna závislosť, t.j. keď jeden z javov – príčina – podmieňuje alebo nevyhnutne vyvoláva za určitých podmienok existenciu druhého javu – účinok (následok). V širšom slova zmysle za príčinnú závislosť považujeme rovnaký priebeh alebo pravidelne sa opakujúcu postupnosť dvoch určitých javov, spôsobenú existenciou tretích príčin, ktoré sú obom javom spoločné.

Skúmanie príčinných závislostí je jedna z najdôležitejších úloh vedy. So skúmaním a vyjadrovaním závislosti medzi dvoma alebo viacerými meniacimi sa veličinami sa stretávame takmer v každej vednej disciplíne.

Príčinná závislosť sa spravidla prejavuje v rôznych formách. Jednou z foriem príčinnej závislosti je funkčná závislosť, s ktorou sa stretávame v matematickej analýze. Podstata funkčnej závislosti je v tom, že každej hodnote jednej premennej veličiny, ktorá sa matematicky nazýva argument a označuje sa zvyčajne písmenom x (veličina nezávisle premenná), sa priradzuje vždy určitá hodnota druhej premennej veličiny, ktorá sa nazýva funkciou a zvyčajne sa označuje písmenom y (závisle premenná veličina). Funkčná závislosť je charakteristická predovšetkým pre prírodné javy – najmä pri niektorých javoch neživej prírody – kým pri spoločenských javoch sa zvyčajne stretávame s veľmi zložitými formami príčinných závislostí. O funkčnej závislosti hovoríme ako aj o pevnej závislosti, pretože vznik alebo existencia jedného javu je pri funkčnej závislosti nerozlučne spojená so vznikom alebo existenciou javu druhého.

Pri skúmaní technologických javov sa spravidla s pevnou závislosťou nestretávame. Tu ide často o spojenie celého komplexu príčin a účinkov. Okrem podstatných vstupných (nezávisle premenných – faktory) pôsobí tu aj celý rad náhodných okolností, náhodných faktorov, pri ktorých v danom okamihu nemožno stanoviť smer, silu ani miesto pôsobenia. Keďže nemôžeme spravidla odstrániť vplyv všetkých vedľajších činiteľov, ani zabezpečiť prísne rovnaké podmienky, môže dôjsť že práve tieto vedľajšie činitele sa tak výhodne, resp. nevýhodne zoskupia, že hodnota sledovaného znaku sa odkláňa od očakávanej úrovne. Ak by sme napr. sledovali pri súbore robotníkov závislosť hrubej mesačnej mzdy od počtu odpracovaných hodín, zistili by sme, že rovnakému počtu odpracovaných hodín u rôznych robotníkov zodpovedá rôzna mesačná mzda; na ich úroveň má totiž vplyv nielen počet odpracovaných hodín, ale aj mnohé ďalšie faktory, o ktorých jednak vieme, jednak nevieme. Určitej hodnote jednej veličiny teda zodpovedá niekoľko hodnôt druhej veličiny, alebo inak povedané, rozloženie početnosti hodnôt druhej veličiny. O týchto veličinách nemôžeme povedať, že je medzi nimi funkčná závislosť. V takomto prípade hovoríme o štatistickej závislosti medzi veličinami. Štatistickú závislosť nazývame aj závislosťou voľnou.

2 Faktory, parametre a objekty výskumu

Stroje, nástroje, prípravky, prístroje a technologické procesy jednotlivo alebo vo vhodnom zoskupení sú objektami výskumu. Vo všeobecnosti sú zložité a mnohotvárne. Typickým príkladom je systém stroj – nástroj – obrobok – prípravok (SNOP). Ich podstata je buď konštantná, alebo sa mení v závislosti od priestoru a času. Stav alebo činnosť strojárskych objektov a technologických procesov charakterizujú dve veličiny. Sú to faktory a parametre.

Na obr.1 je kybernetický model („čierna skrinka“) objektu výskumu so vstupmi aj výstupmi.

Faktory (nezávisle premenné) sú vstupné veličiny objektu výskumu a určujú jeho správanie sa. Rozdeľujú sa na :

riadené faktory x_j , $j = 1, 2, \dots, m$; ktoré majú pevnú hodnotu alebo sa menia počas pokusu vopred zadaným spôsobom (napr. konštrukčné a geometrické prvky nástroja, rezné podmienky, množstvo chladiacej kvapaliny a pod.)

kontrolované (neriadené) faktory u_k , $k = 1, 2, \dots, m$. Ich hodnoty sa pri celom pokuse kontrolujú a udržiavajú na konštantnej hodnote. Nie sú predmetom pokusu.

Rušiace (neriadené a nekontrolované; šumové) w_p , $p = 1, 2, \dots, m$. Ide o faktory neznámeho pôvodu, s neznámou fyzikálnou podstatou, neurčeného množstva a druhu, s neznámym miestom a silou pôsobenia (napr. zmena tvrdosti v rôznych miestach polovýrobku).

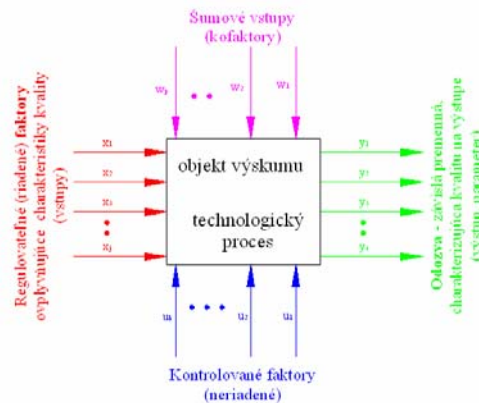
Parametre (optimalizačné parametre, cieľová funkcia, optimalizačné kritérium) y_i , $i = 1, 2, \dots, n$, sú výstupné veličiny objektu výskumu a vyjadrujú jeho zmenu vplyvom jednotlivých faktorov (opotrebenie nástroja, rezné sily, teplota rezania, drsnosť obrobených plôch, presnosť obrábania...).

Objekty výskumu možno klasifikovať ako preskúmané, nepreskúmané, riadené a neriadené, jednofaktorové a viacfaktorové, jednoparametrové a viacparametrové, statické a dynamické, stacionárne a nestacionárne, atď.

Aktívny experiment si vyžaduje, aby objekt výskumu bol riadený, a aby výsledky pokusov boli reprodukovateľné.

Závislosť medzi parametrami a faktormi nazývame matematický model (funkcia reakcií) :

$$\eta_i = \varphi(x_1, x_2, \dots, x_m), \quad i = 1, 2, \dots, n \quad (1.1)$$



Obr.1 Objekt výskumu

3 Regresná analýza

Korelačná a regresná analýza má základy v práci Francisa Galtona (1894). Čiastočne inšpirovaný dielom svojho bratranca Charlesa Darwina, O pôvode druhov, snažil sa Galton odhaliť dedičné vlastnosti talentu, pohybových schopností a intelektu. Jeho výskum začal štúdiom veľkosti a váhy bieleho hrachu v dvoch generáciách. Ten viedol k odhaleniu fenoménu regresie, ktorú Galton nazval „reverse“ (návrat). Galton zistil, že potomkovia extrémne veľkých hrachov neboli v priemere taký veľký, ako ich „rodičia“. Neskôr zaznamenával fyzické charakteristiky tisícov dobrovoľníkov a našiel podobnú „regresiu“ k priemeru pri medzigeneračnom porovnávaní. Jeho koncepty potom rozvinul jeho žiak K. Pearson a ďalší štatistickí. Tradičný názov „regresia“ sa stále používa v dôsledku tradície, hoci sa táto technika používa predovšetkým pre predikciu. Metódu najmenších štvorcov pre odhad regresných koeficientov regresných funkcií používali už A.M. Legendre a C.F. Gauss.

Hlavnou úlohou regresnej a korelačnej analýzy je prispieť k poznaniu príčinných vzťahov medzi štatistickými znakmi. Východiskom pre popis štatistických závislostí sú štatistické údaje. Štatistický súbor n pozorovaní sledovaní štatistických znakov možno získať niekoľkými spôsobmi :

- pozorovaním n štatistických jednotiek, pričom základný súbor bol priestorovo, časovo aj vecne vymedzený,
- pozorovaním určitej štatistickej jednotky v n časových okamihoch alebo intervaloch,
- násobným opakovaním určitého pokusu, realizovaného pri rovnakých, resp. približne rovnakých podmienok.

Úlohou regresnej, resp. korelačnej analýzy je ďalej matematický popis systematických okolností, ktoré sprevádzajú štatistické závislosti.

Pri budovaní regresných modelov sa bežne používa metóda najmenších štvorcov (ďalej MNŠ). Táto metóda poskytuje postačujúce odhady parametrov iba pri súčasnom splnení všetkých predpokladov o dátach a o regresnom modeli. Pokiaľ tieto predpoklady nie sú splnené strácajú výsledky MNŠ svoje vlastnosti. Využitie lineárnej regresnej analýzy sa dotýka nasledujúcich možností :

1. Popis dát : hľadáme vzťah, lineárny regresný model, ktorý sumarizuje súbor dát

2. Určenie parametrov : najbežnejším cieľom regresnej analýzy je vyčíslenie najlepších odhadov neznámych parametrov regresného modelu. Užívateľ navrhne regresný model a regresnou analýzou sa snaží model preukázať. Tento cieľ často pokrýva aj ostatné zámery regresnej analýzy.
3. Predikcia : najdôležitejším cieľom regresnej analýzy je predikcia, vyčíslenie hodnôt závisle premennej. Je tu však rad predpokladov a kvalifikácií, ktoré sa musia rešpektovať v regresnom modeli a dátach. Často sa napríklad nesmie extrapolovať mimo rozsahu dát. Intervalové odhady vyžadujú dodržanie predpokladov normality. MNŠ má svojich sedem predpokladov, ktoré sa musia rešpektovať a dodržať.
4. Riadenie : regresné modely možno využiť tiež na monitorovanie a riadenie systémov, ako napríklad kalibrácii meracieho systému. Ak sa využije regresný model na riadiace účely, nezávisle premenné musia byť vzťahované k závisle premennej kauzálnym spôsobom.
5. Výber (voľba) premenných : voľba premenných sleduje tie nezávisle premenné, ktoré vysvetľujú významný objem premenlivosti závisle premennej. V množstve aplikácií sa nejedná o jednorazový proces, ale o spojitý proces výstavby modelu.

V nasledujúcej časti sa pokúsime matematicky formulovať regresný problém ako aj vysvetliť odhad regresných koeficientov MNŠ.

Uvažujme základný štatistický súbor, v ktorom skúmame štatistické znaky $y, x_1, x_2, \dots, x_k$. Ak sa mení nejakým spôsobom podmienené rozdelenie znaku y pri zmenách znakov $x_1, x_2, \dots, x_k$, hovoríme o štatistickej závislosti znaku y na znakoch $x_1, x_2, \dots, x_k$. Špeciálnym prípadom štatistickej závislosti je korelačná závislosť, keď sa menia podmienené stredné hodnoty η_i znaku y . Znak y nazývame vysvetľovanou premennou, alebo závisle premennou a znaky $x_1, x_2, \dots, x_k$ vysvetľujúce resp. nezávisle premenné. Regresný model

$$y_i = \eta_i + \varepsilon_i \quad (1.2)$$

vyjadruje i – tú hodnotu y_i závisle premennej ako súčet podmienenej strednej hodnoty η_i závisle premennej y pri kombinácii hodnôt nezávisle premenných $x_{1i}, x_{2i}, \dots, x_{ki}$ a náhodnej zložky ε_i . Podmienenú strednú hodnotu η ako funkciu nezávisle premenných nazývame regresnou funkciou.

Podľa tvaru regresnej funkcie rozlišujeme rôzne typy regresných modelov :

- modely lineárne z hľadiska parametrov majú regresnú funkciu v tvare :

$$\eta = \beta_0 + \beta_1 f_1 + \beta_2 f_2 + \dots + \beta_r f_r \quad (1.3)$$

kde regresory $f_1, f_2, \dots, f_r$ sú ľubovoľné známe funkcie nezávisle premenných $x_1, x_2, \dots, x_k$.

- Špeciálnym prípadom sú modely s regresnou funkciou :

$$\eta = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k \quad (1.4)$$

kde regresory sú priamo nezávisle premenné, tj. modely lineárne z hľadiska parametrov aj z hľadiska nezávisle premenných.

4 Metóda najmenších štvorcov

Všeobecne používaná MNŠ pri tvorbe regresných modelov poskytuje postačujúce odhady parametrov regresnej funkcie iba pri splnení nasledovných predpokladov :

1. Regresné parametre β môžu nadobúdať ľubovoľné hodnoty. V praxi však často existujú obmedzenia faktorov a parametrov, ktoré vychádzajú z ich fyzikálneho zmyslu.
2. Regresný model je lineárny v parametroch a platí aditívny model merania.
3. Matica nenáhodných, nastavovaných hodnôt vysvetľujúcich premenných $\mathbf{X}$ má hodnotu práve m , čo znamená, že žiadne ich dva stĺpce x_j, x_k nie sú kolineárne, t.j. rovnobežné vektory. Tomu odpovedá aj formulácia, že matica $\mathbf{X}^T \mathbf{X}$ je symetrická regulárna matica, ku ktorej existuje inverzná matica a determinant matice $\mathbf{X}^T \mathbf{X}$ je väčší ako nula.
4. Náhodné chyby ε_i majú nulovú strednú hodnotu $E(\varepsilon_i)=0$. To musí pri korelačných modeloch platiť vždy.
5. Náhodné chyby ε_i majú konštantný a konečný rozptyl. Tiež podmienený rozptyl $D(y/x) = \sigma^2$ je konštantný a ide o homoskedastický prípad.
6. Náhodné chyby ε_i sú navzájom nekorelované a platí $\text{cov}(\varepsilon_i, \varepsilon_i) = E(\varepsilon_i \varepsilon_i) = 0$. Pokiaľ majú chyby normálne rozdelenie, sú nezávislé. Táto požiadavka odpovedá požiadavke nezávislosti meranej veličiny y .
7. Chyby ε_i majú normálne rozdelenie $N(0, \sigma)$. Vektor y má potom navyše viacrozmerné normálne rozdelenie so strednou hodnotou $\mathbf{X}\beta$ a kovariačnou maticou $\sigma^2 \mathbf{E}$, kde $\mathbf{E}$ je jednotková matica.

Pokiaľ platia predpoklady 1. až 6., sú odhady $\mathbf{b}$ parametrov β najlepšie, nestranné a lineárne. Navyše majú asymptoticky normálne rozdelenie. Pokiaľ platí ešte aj predpoklad 7. majú odhady $\mathbf{b}$ normálne rozdelenie aj pre konečné výbery.

Výpočet koeficientov regresnej funkcie môžeme realizovať aj pomocou maticových operácií. Si V ďalšej časti je uvedený principiálny výpočet regresnej závislosti v maticovom tvare. Regresná funkcia v maticovom tvare :

$$\mathbf{Y} = \mathbf{X} \cdot \beta + \varepsilon \quad (1.5)$$

$\mathbf{Y}$ – stĺpcový vektor výsledkov meraní ($y_i, i=1,2,...,n$)

$\mathbf{X}$ – matica podmienok merania, pričom i – tému výsledku merania y_i odpovedá i - tá podmienka merania (v_{ci}, f_i, r_{ei})

β – stĺpcový vektor regresných koeficientov

ε – vektor náhodných chýb

Rovnicu (1.5) môžeme rozpísať ako :

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_i \end{bmatrix} = \begin{bmatrix} x_{01} & x_{11} & \cdot & \cdot & x_{n1} \\ x_{02} & x_{12} & \cdot & \cdot & x_{n2} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ x_{0i} & x_{1i} & \cdot & \cdot & x_{ni} \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \cdot \\ \beta_k \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \cdot \\ \varepsilon_i \end{bmatrix} \quad (1.6)$$

kde x_{0i} predstavuje fiktívnu premennú ($x_0 = 1$ pre lineárnu resp. $x_0 = 10$ pre logaritmickú závislosť)

$$\boldsymbol{\beta} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_k \end{bmatrix} = [\mathbf{X}^T \cdot \mathbf{X}]^{-1} \cdot \mathbf{X}^T \cdot \mathbf{Y} \quad (1.7)$$

kde $\mathbf{X}^T$ – transponovaná matica k matici $\mathbf{X}$

$[\mathbf{X}^T \cdot \mathbf{X}]^{-1}$ – inverzná matica súčinu transponovanej matice podmienok merania $\mathbf{X}$ a matice $\mathbf{Y}$.

$[\mathbf{X}^T \cdot \mathbf{X}]$ – Fisherova informačná matica

Aby vo vzťahu (1.7) existovala inverzná matica $[\mathbf{X}^T \cdot \mathbf{X}]^{-1}$ je potrebné, aby matica $[\mathbf{X}^T \cdot \mathbf{X}]$ bola regulárna, teda jej determinant $|\mathbf{X}^T \cdot \mathbf{X}| \neq 0$ (v opačnom prípade je matica singulárna). Z toho vyplýva, že na použitie uvedenej metódy je nevyhnutné, aby nezávislé premenné ($x_1, x_2, \dots, x_i$) boli lineárne nezávislé. V takomto prípade v matici nezávisle premenných $\mathbf{X}$ prvky ľubovoľného stĺpca nebudú lineárnou kombináciou odpovedajúcich prvkov druhých stĺpcov.

Symetrická štvorcová matica $[\mathbf{X}^T \cdot \mathbf{X}]^{-1}$ sa nazýva kovariačná matica, pre ktorú platí :

$$E[\mathbf{b} - \boldsymbol{\beta}]^T \cdot [\mathbf{b} - \boldsymbol{\beta}] = [\mathbf{X}^T \cdot \mathbf{X}]^{-1} \cdot s_r^2 = \text{VAR}[\mathbf{b}] =$$

$$= \begin{bmatrix} s_{b0}^2 & \text{cov}(b_0, b_1) & \cdot & \cdot & \text{cov}(b_0, b_k) \\ \text{cov}(b_1, b_0) & s_{b1}^2 & \cdot & \cdot & \text{cov}(b_1, b_k) \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \text{cov}(b_k, b_0) & \text{cov}(b_k, b_1) & \cdot & \cdot & s_{bk}^2 \end{bmatrix} \quad (1.8)$$

kde : s_r^2 – reziduálny rozptyl

s_{bi}^2 – celkový reziduálny rozptyl regresného koeficientu b_i

$\text{cov}(b_i, b_k)$ – kovariancia alebo korelačný pomer koeficientov b_i a b_k

$\mathbf{E}$ – jednotková matica

$\boldsymbol{\beta}$ – stĺpcový vektor koeficientov skutočnej regresie

Prvky hlavnej diagonály matice (1.8) $s_{b1}^2, s_{b2}^2, \dots, s_{bi}^2$ predstavujú rozptyl koeficientov regresnej funkcie, ktoré sú potrebné na testovanie ich štatistickej významnosti. Nediagonálne prvky sú kovariancie odpovedajúcich regresných koeficientov, vyjadrujúce štatistickú závislosť regresných koeficientov regresnej rovnice (kovariancia je miera závislosti dvoch

premenných $\text{cov } xy = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})$, teda je to priemerný súčin odchýlok

korelovaných hodnôt od ich aritmetických priemerov).

Následne je potrebné urobiť kontrolu správnosti inverzie súčinu matíc podľa vzťahu :

$$\mathbf{E} = [\mathbf{X}^T \cdot \mathbf{X}]^{-1} \cdot [\mathbf{X}^T \cdot \mathbf{X}] \quad (1.9)$$

Matica **E** predstavuje jednotkovú maticu, s jednotkami na hlavnej diagonále a nulovými ostatnými prvkami. Určité odchýlky od nulovej hodnoty môžu byť spôsobené chybami pri zaokrúhľovaní a pri použití linearizácie.

Pretože pre maticový súčin platí asociatívnosť môže sa táto skutočnosť využiť pre dosiahnutie medzivýsledku

$$\mathbf{A} = [\mathbf{X}^T \cdot \mathbf{X}]^{-1} \cdot \mathbf{X}^T \quad (1.10)$$

Vektor odhadov koeficientov regresnej funkcie je potom :

$$\mathbf{b} = \mathbf{A} \cdot \mathbf{Y} \quad (1.11)$$

Empirická hodnota y je hodnota, ktorá je získaná ako výsledok merania a teoretická hodnota $\hat{y}$, ktorá je získaná ako výsledok modelovania tejto premennej. Rozdiel medzi teoretickou a empirickou hodnotou je označený ako rezíduum e :

$$e = y_i - \hat{y}_i \quad (1.12)$$

čo je veľkosť chyby v príslušnom bode.

V nasledujúcej časti je uvedený skript vytvorený vo výpočtovom systéme MATLAB, ktorý bol aplikovaný na vytvorenie matematicko – štatistického modelu pri skúmaní závislosti strednej aritmetickej odchýlky profilu drsnosti v závislosti od rezných podmienok pri pozdĺžnom sústružení ocele 12 050.1, ako aj pri vytvorení matematického modelu závislosti makrogeometrie a mikrogeometrie povrchu v závislosti od pracovných podmienok pri rezaní nehrdzavejúcej ocele a hliníka vysokorýchlostným hydroabrazívnym prúdom. Jedná sa však iba o tú časť, pomerne rozsiahleho skriptu, v ktorej je riešený práve výpočet regresných koeficientov modelu.

```
%skript pre vyhodnotenie meraní pomocou maticovej regresie
%určený pre mocninovu regresnu funkciu - planovany experiment
%log y=b0*log(10)+b1*log(v_c)+b2*log(f)+b3*log(a_p)
%y=10^b0*(v_c^b1)*(f^b2)*(a_p^b3)
%y=Ra, Rz
clc;
format long
load matica.txt; % načítanie výsledkov meraní z matice
matica %zobrazenie matice meraní a podmienok experimentu
[poc_pripadov,poc_stlpcov]=size(matica)
N=poc_pripadov;
load FISHER_095.txt;%načítanie matice kvantilov Fisherovho rozdelenia
load STUDENT_0975.txt;%načítanie matice kvantilov Studentovho rozdelenia
load CHI_095.txt;%načítanie matice kvantilov Chi rozdelenia
load meranie.txt % načítanie výsledkov merania z matice meranie
meranie=log10(meranie) %zobrazenie matice meranie
[pripady,stlpc]=size(meranie)%stanovenie počtu prípadov N a opakovaných
meraní m
m=stlpc%počet opakovateľnosti meraní
M=log10(matica);%všetky prvky matice sa prevedú na dekadické logaritmy
y=M(:,5); %stĺpec parametra merania do premennej y
x0=M(:,1);%načítanie prvého stĺpca-fiktívna premenná do premennej x0
x1=M(:,2);%načítanie druhého stĺpca do premennej x1
x2=M(:,3);%načítanie tretieho stĺpca do premennej x2
x3=M(:,4);%načítanie štvrtého stĺpca do premennej x3
Y=y; %vytvorenie stĺpcového vektora Y
X=[x0 x1 x2 x3]; % vytvorenie matice X
Xt=X';%transponovaná matica k matici X
Xi=Xt*X %súčin transponovanej matice X a matice X
DXi=det(Xi) %determinant matice Xi
```

```

Xinv=inv(Xi) %inverzná matica k matici Xi
H=Xinv;% Fisherova informačná matica
E=Xinv*Xi %jednotková matica-kontrola správnosti inverzie
A=Xinv*Xt; %pomocná matica pre výpočet koeficientov regresie
b=A*Y %vektor koeficientov regresnej funkcie

```

Záver

Predložený príspevok sa zaoberá princípom tvorby regresného modelu. Vytvoriť regresný model, určiť odhady jeho regresných koeficientov, je však iba jedným z prvých krokov pri výstavbe modelu technologického procesu alebo systému, pretože následne je potrebné analyzovať tzv. regresný triplet (dáta, metóda, model) aby nami vytvorený štatisticko – matematický model bol naozaj správny a funkčný. Spojenie znalostí štatistických metód analýzy závislostí, testovania štatistických hypotéz, výrobných technológií a metód plánovania experimentov (DOE) dokáže byť veľmi silným nástrojom v reálnej prevádzke výrobných podnikov pri jej optimalizácii v spojení s modernou výpočtovou technikou. Na tomto mieste je nutné ale podotknúť, že žiaden výpočtový systém nedokáže vyriešiť problém za nás. Počítač a príslušný výpočtový software je iba nástrojom, pomocníkom ktorý uľahčuje výpočty. Otázky čo chcem vypočítať ? ako to chcem vypočítať ? a ako si overím výsledok? za nás žiaden „program“ nevyrieši.

Zoznam odkazov a informačné zdroje

1. KAŽIMÍR, I., BEŇO, J. : *Teória obrábania – Návod na cvičenia*. Bratislava : ALFA, 1989, 280 s.
2. LINCZENYI, A.: *Inžinierska štatistika*. Bratislava : ALFA, 1974, 456 s.
3. MELOUN, M., MILITKÝ, J. : *Kompendium statistického zpracování dat*. Praha : ACADEMIA, 2001, 766 s., ISBN : 80-200-1008-4
4. MILITKÝ, J.: *Matlab a analýza dat*. Liberec : Katedra textilních materiálu TU v Liberci, 2002, 149 s. Dostupné na internete:
<http://www.ft.vslib.cz/study/akreditace/akreditace2004/priloha8/programovanivmatlabu/matprir.pdf>

Lektoroval: prof. Ing. Emil Ragan, CSc.

Kontakt

Ing. Miroslav Gombár, PhD., Ústav digitálnych kompetencií PU, 17. Novembra 11, 080 01 Prešov, mirek@unipo.sk

Ing. Sergej Hloch, PhD., Katedra prevádzky technologických systémov FVT TU Košice so sídlom v Prešove, Štúrová 13, 080 01 Prešov, hloch.sergej@fvt.sk

Ing. Ján Valíček, Ph.D., Institut fyziky, Hornicko geologická fakulta, VŠB – Technická univerzita, Ostrava, Tř. 17 listopadu 15/2171, 708 33 Ostrava – Poruba, Česká republika, jan.valicek@vsb.cz

Ing. Alexander Pella, Ústav digitálnych kompetencií PU, 17. Novembra 11, 080 01 Prešov, mirek@unipo.sk