Liet. mat. rink. LMD darbai, **50**, 2009, 120–124

Eksperimento planavimo dėstymas naudojant kompiuterius

Vidmantas Povilas PEKARSKAS
Kauno technikos universitetas
Studentų g. 50, LT-51368 Kaunas
el. paštas: vidmantas.pekarskas@ktu.lt

Santrauka. Straipsnyje nagrinėjamas konkretus informacinių technologijų taikymo pavyzdys, dėstant kursą "Eksperimento planavimas". Parodyta, kaip panaudojant kompiuterius, galima ne tik spręsti eksperimento planavimo uždavinius, iliustruoti teoriją, bet ir pagrįsti kai kuriuos teorinius teiginius.

Raktiniai žodžiai: eksperimento planavimas, paketo Mathcad taikymas.

Šio straipsnio **tikslas** – pasidalyti patirtimi, kaip galima dėstyti eksperimento planavimo kursą, pasitelkiant matematinį paketą Mathcad bei išryškinti tokios kurso dėstymo metodikos privalumus, lyginant su tradiciniu kurso dėstymu.

Eksperimento planavimo kursas dėstomas KTU Dizaino ir technologijų fakulteto aprangos ir polimerinių gaminių bei medienos technologijos studijų programos magistrantams bei Fundamentaliųjų mokslų fakulteto taikomosios fizikos studijų programos magistrantams. Šį kursą išklauso ir įvairių specialybių KTU doktorantai, kurie mano, kad jis bus jiems naudingas, planuojant eksperimentą ir apdorojant jo rezultatus. Eksperimento planavimo kursą paprastai klauso apie 20 studentų. Užsiėmimai vyksta kompiuterių klasėje.

Informacinių technologijų diegimo procesas edukologijoje skirstomas į kelis etapus [2]. Šiame straipsnyje aprašytą informacinių technologijų taikymo modelį galima priskirti prie būdų, kuriais ne tik įvairinamas tradicinis edukacinis procesas ir didinamas jo veiksmingumas, bet ir išplečiamos mokymo ir mokymosi galimybės.

Teorinė kurso dalis išdėstoma tradiciškai kaip paskaita. Uždaviniai sprendžiami pasitelkiant kompiuterius bei paketą Mathcad [1]. Eksperimentas planuojamas priklausomai nuo kelių faktorių $x_1, x_2, \ldots, x_k$, o eksperimento rezultatai paprastai aproksimuojami pirmojo arba antrojo laipsnio kelių kintamųjų daugianariais. Kai faktorių reikšmes, su kuriomis atliekamas eksperimentas, apibūdina matrica

$$X = \begin{pmatrix} x_{01} & x_{11} & x_{21} & \ldots & x_{k1} \\ x_{02} & x_{12} & x_{22} & \ldots & x_{k2} \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ x_{0N} & x_{1N} & x_{2N} & \ldots & x_{kN} \end{pmatrix},$$

o gautus eksperimento rezultatus – matrica $Y = (y_1\ y_2\ \ldots\ y_N)^T$, tai modelio

$$y = b_0 + b_1 x_1 + b_2 x_2 + \ldots + b_k x_k \tag{1}$$

koeficientus apibrėžianti matrica $B = (b_0, b_1, \ldots, b_k)^T$, gaunama iš sąlygos

$$B = (X^T X)^{-1} X^T Y. \tag{2}$$

Modelio koeficientų $b_0, b_1, \ldots, b_k$ apskaičiavimo formulės priklauso nuo eksperimento plano, apibūdinamo matrica X, charakterio. Koeficientų išraiška ypač supaprastėja, kai planavimas yra ortogonalusis. Taip vadinamas planas, kai bet kurių dviejų to plano matricos X stulpelių skaliarinės sandaugos yra lygios nuliui. Norint gauti kvadratinio modelio koeficientus, naudojamasi ortogonaliais centriniais kompoziciniais arba rototabeliaisiais planais. Tokių planų matrica X sudaroma pagal tam tikras taisykles, o planai vėlgi pasižymi tik jiems būdingomis savybėmis.

Eksperimento planavimo kurse išvedamos formulės, skirtos (1) regresijos lygties koeficientams apskaičiuoti. Šios formulės ypač paprastos, kai eksperimento planas yra pilnasis faktorinis. Kitais atvejais formulės yra žymiai sudėtingesnės, todėl apskaičiuoti koeficientus, panaudojant pavyzdžiui Excel, nėra taip lengva. Tai pailiustruoja žemiau pateikiamos formulės.

Štai norint gauti kvadratinio modelio

$$\begin{aligned} y = {} & b_0 + b_1 x_1 + b_2 x_2 + \ldots + b_k x_k + b_{12} x_1 x_2 + b_{13} x_1 x_3 + \ldots + b_{k-1} x_{k-1} x_k \\ & + b_{11} x_1^2 + b_{22} x_2^2 + \ldots + b_{kk} x_k^2 \end{aligned} \tag{3}$$

koeficientus, kai planas yra rototabelusis, naudojamos tokios jų formulės:

$$\begin{aligned} b_0 &= a_1 \sum_{u=1}^{N} y_u + a_2 \sum_{i=1}^{k} \sum_{u=1}^{N} x_{iu}^2 y_u, \\ b_i &= a_3 \sum_{u=1}^{N} x_{iu} y_u, \quad b_{ij} = a_4 \sum_{u=1}^{N} x_{iu} x_{ju} y_u, \\ b_{ii} &= a_5 \sum_{u=1}^{N} x_{iu}^2 y_u + a_6 \sum_{i=1}^{k} \sum_{u=1}^{N} x_{iu}^2 y_u + a_7 \sum_{u=1}^{N} y_u; \end{aligned} \tag{4}$$

čia a_i $(i = \overline{1,7})$ tam tikri koeficientai, kurių reikšmės priklauso nuo faktorių skaičiaus k.

Pasitelkę Mathcad, visais atvejais galime (3) regresijos lygties koeficientus apskaičiuoti tiesiogiai pagal (2) formulę, nenaudodami tiems koeficientams apskaičiuoti skirtų sudėtingų (4) formulių. Svarbu tik tai, kad matricą sudarytų ne daigiau kaip 600 elementų (toks yra Mathcad 14 apribojimas). Taigi, taikydami (2) formulę, galime apskaičiuoti (3) modelio koeficientus, kai ortogonalusis centrinis kompozicinis arba rototabelusis planas yra sudarytas dviems, trims arba keturiems faktoriams. Kai faktorių yra penki, tai, pavyzdžiui, ortogonaliojo centrinio kompozicinio plano matricą sudaro 903 elementai ir taikyti (2) formulės jau negalima. Tai nėra didelis trūkumas, nes praktikoje retai planuojami eksperimentai, parenkant faktorių skaičių $k \geqslant 4$. Mathcad taikymo pavyzdį pateiksime apdorodami eksperimento duomenis, kurie buvo gauti realizavus rototabelųjį planą, kai faktorių yra trys. Šiame pavyzdyje panaudoti realiojo eksperimento rezultatai, gauti tiriant sanklijų stiprį y, priklausomai nuo tam tikrų cheminių komponentų x_1, x_2, x_3, įterptų į klijų sudėtį. Rototabelųjį planą apibūdina matrica X, gautas eksperimento reikšmes – matrica Y, regresijos lygties

$$y = b_0 + b_1x_1 + b_2x_2 + b_3x_3 + b_{11}x_1^2 + b_{22}x_2^2 + b_{33}x_3^2 + b_{12}x_1x_2 \\ + b_{13}x_1x_3 + b_{13}x_2x_3$$

koeficientus – matrica $(X^T X)^{-1} X^T Y$. Žemiau pateiktas uždavinio sprendimas, panaudojant paketą Mathcad.

$$X := \begin{pmatrix} 1 & -1 & -1 & -1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 & 1 & 1 & 1 & -1 & -1 & 1 \\ 1 & -1 & 1 & -1 & 1 & 1 & 1 & -1 & 1 & -1 \\ 1 & 1 & 1 & -1 & 1 & 1 & 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 & 1 & 1 & 1 & 1 & -1 & -1 \\ 1 & 1 & -1 & 1 & 1 & 1 & 1 & -1 & 1 & -1 \\ 1 & -1 & 1 & 1 & 1 & 1 & 1 & -1 & -1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1.682 & 0 & 0 & 2.828 & 0 & 0 & 0 & 0 & 0 \\ 1 & -1.682 & 0 & 0 & 2.828 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1.682 & 0 & 0 & 2.828 & 0 & 0 & 0 & 0 \\ 1 & 0 & -1.682 & 0 & 0 & 2.828 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1.682 & 0 & 0 & 2.828 & 0 & 0 & 0 \\ 1 & 0 & 0 & -1.682 & 0 & 0 & 2.828 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} . Y := \begin{pmatrix} 3.03 \\ 3.83 \\ 2.34 \\ 4.05 \\ 3.78 \\ 3.14 \\ 3.69 \\ 3.97 \\ 3.75 \\ 2.84 \\ 3.89 \\ 3.75 \\ 3.73 \\ 3.18 \\ 4.26 \\ 4.27 \\ 4.25 \\ 4.28 \\ 4.24 \\ 4.21 \end{pmatrix} . (X^T X)^{-1} X^T Y = \begin{pmatrix} 4.252 \\ 0.269 \\ 0.037 \\ 0.165 \\ -0.338 \\ -0.153 \\ -0.282 \\ 0.229 \\ -0.359 \\ 0.151 \end{pmatrix}.$$

Aiškiai matyti, kad taip išspręsti uždavinį daug paprasčiau, negu taikant (4) formules. Reikia pabrėžti ir tai, kad taikydami paketą Mathcad studentai gali ne tik daugiau išspręsti uždavinių, varijuodami planų duomenis, bet ir ištirti duomenų kaitos įtaką gaunamiems rezultatams. Iš to jie gali savarankiškai padaryti tam tikras išvadas apie planų savybes. Toks mokymosi būdas sužadina studentų pažinimą labiau, negu tradicinis edukacinis metodas.

Panaudojant paketą Mathcad, labai patogu tirti modelio adekvatumą. Didaktiniu požiūriu čia svarbu ir tai, kad atsiranda galimybių pademonstruoti, kaip nuo stebėjimų tikslumo bei nuo parinkto reikšmingumo lygmens priklauso eksperimentatoriaus sprendimas apie modelio adekvatumą.

Tarkime, eksperimento planas ir jo rezultatai apibūdinami 1 lentele.

1 lentelė. Eksperimento duomenys

i	x_1	x_2	y_{1u}	y_{2u}	$y_u = \frac{1}{2}(y_{1u} + y_{2u})$
1	0	0	76,8	79,4	78,1
2	1	0	80,5	77,9	79,2
3	0	1	73,4	70,9	72,15
4	–1	0	70,3	71,2	70,75
5	0	–1	76,2	76,3	76,25
6	1	1	72,1	73,2	72,65
7	1	–1	79,0	80,2	79,6
8	–1	–1	72,2	72,1	72,15
9	–1	1	69,1	69,18	69,14

Rasime modelio $y = b_0 + b_1 x_1 + b_2 x_2 + b_{11} x_1^2 + b_{22} x_2^2 + b_{12} x_1 x_2$ koeficientus ir patikrinsime modelio adekvatumą, panaudodami statistiką $F = \frac{s^2_{nead}}{s^2\{y\}}$, kuri turi Fišerio skirstinį su φ_1 ir φ_2 laisvės laipsnių. Uždavinio sprendimas, panaudojant paketą Mathcad pateikiamas žemiau. Neadekvatumo dispersijos įvertis s^2_{nead} toliau dėl paprastumo pažymėtas $s1$, vienetinio stebinio paklaidos dispersijos įvertis $s^2\{y\} - s2$.

$$X := \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & -1 & 0 & 1 & 0 & 0 \\ 1 & 0 & -1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & 1 & 1 & -1 \\ 1 & -1 & -1 & 1 & 1 & 1 \\ 1 & -1 & 1 & 1 & 1 & -1 \end{pmatrix}, \quad Y := \begin{pmatrix} 76.8 & 79.4 & 78.1 \\ 80.5 & 77.9 & 79.2 \\ 73.4 & 70.9 & 72.15 \\ 70.3 & 71.2 & 70.75 \\ 76.2 & 76.3 & 76.25 \\ 72.1 & 73.2 & 72.65 \\ 79.0 & 80.2 & 79.6 \\ 72.2 & 72.1 & 72.15 \\ 69.1 & 69.18 & 69.14 \end{pmatrix}, \quad (X^T X)^{-1} X^T Y^{\langle 2 \rangle} = \begin{pmatrix} 77.073 \\ 3.235 \\ -2.343 \\ -1.585 \\ -2.36 \\ -0.985 \end{pmatrix},$$

$$y := 77.073 + 3.235 X^{\langle 1 \rangle} - 2.343 X^{\langle 2 \rangle} - 1.585 X^{\langle 3 \rangle} - 2.36 X^{\langle 4 \rangle} - 0.985 X^{\langle 5 \rangle}.$$

$$Y = \begin{pmatrix} 77.073 \\ 78.723 \\ 72.37 \\ 72.253 \\ 77.056 \\ 73.035 \\ 79.691 \\ 71.251 \\ 68.535 \end{pmatrix}, \quad SR := 2 \sum_{i=0}^{8} \left[\left(Y^{\langle 2 \rangle} - y \right)^2 \right]_i, \quad SR = 11,14, \quad \Phi 1 := 3,$$

$$s1 := \frac{SR}{\Phi 1}, \quad s1 = 3.713,$$

$$SE := \sum_{i=0}^{8} \left[\left[\left(Y^{\langle 2 \rangle} - Y^{\langle 0 \rangle} \right)^2 \right]_i + \left[\left(Y^{\langle 2 \rangle} - Y^{\langle 1 \rangle} \right)^2 \right]_i \right], \quad SE = 11.628, \quad \Phi 2 := 9,$$

$$s2 := \frac{SE}{\Phi 2}, \quad s2 = 1.292, \quad \underset{\sim}{F} := \frac{s1}{s2}, \quad F = 2.874$$

$$qF(0.95;\ 3;\ 9) = 3.863.$$

Kadangi $F < qF(0,95;\ 3;\ 9)$, tai modelis adekvatus. Nesunku įsitikinti, kad esant tam pačiam reikšmingumo lygmeniui $\alpha = 0,05$, modelis tampa neadekvatus, jei tik padidinamas stebėjimų tikslumas. Pakanka patikslinti pirmąsias tris y_{1u} ir y_{2u} reikšmių poras, parenkant jas artimesnes vidurkiui y_u, pavyzdžiui,

$$y_{11} = 77,8, \ \ y_{21} = 78,4; \ \ y_{12} = 79,5, \ \ y_{22} = 78,9 \text{ ir} y_{13} = 72,4, \ \ y_{23} = 71,9.$$

Tuomet gaunama

$$SE = 2,228, \quad \Phi 2 := 9, \quad s2 := \frac{SE}{\Phi 2}, \quad s2 = 0,248,$$

$$\underset{\sim}{F} := \frac{s1}{s2}, \quad F = 14,999, \quad qF(0,95;\ 3;\ 9) = 3,863.$$

Tai reiškia, kad hipotezė apie modelio adekvatumą atmetama. Taip keisdami eksperimento duomenis bei reikšmingumo lygmenį α nesunkiai galime pademonstruoti, kas gali įtakoti eksperimentatoriaus sprendimą apie modelio adekvatumą. Šitaip ne tik pademonstruojama uždavinio apie modelio adekvatumą sprendimo metodika, bet ir pagilinamas studento suvokimas apie statistinių hipotezių tikrinimą.

Eksperimento planavime svarbus vaidmuo tenka matricai $(X^T X)^{-1}$, kuri vadinama koreliacine matrica. Panaudojant ją, apskaičiuojamos regresijos lygties koeficientų dispersijų bei koeficientų kovariacijų įverčiai, taip pat išsiaiškinama, kurie koeficientai nustatomi nepriklausomai vienas nuo kito. Kai matrica X turi daug elementų, tai rasti matricą $(X^T X)^{-1}$, naudojantis žinoma atvirkštinės matricos formule, sudėtinga. Čia labai padeda paketas Mathcad. Ankstesniam pavyzdžiui būdinga tokia matrica $(X^T X)^{-1}$:

$$(X^T X)^{-1} = \begin{pmatrix} 0.556 & 0 & 0 & -0.333 & -0.333 & 0 \\ 0 & 0.167 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.167 & 0 & 0 & 0 \\ -0.333 & 0 & 0 & 0.5 & 0 & 0 \\ -0.333 & 0 & 0 & 0 & 0.5 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0.25 \end{pmatrix}$$

Iš jos matyti, kad planas beveik ortogonalus ir dauguma regresijos lygties koeficientų nustatoma nepriklausomai vienas nuo kito.

Panaudojant paketą Mathcad galima iliustruoti ir daugiau teorinių teiginių. Pavyzdžiui, teiginį apie tai, kad neįmanoma gauti (3) modelio koeficientų, kai visi faktoriai varijuojami tik dviejuose lygiuose +1 ir – 1. Norint tuo įsitikinti, pakanka apskaičiuoti matricos $X^T X$ determinantą, kuris gaunamas lygus nuliui.

Kelių metų pedagoginė patirtis, įgyta taip dėstant eksperimento planavimo kursą, leidžia padaryti tokias **išvadas**:

- pavyksta optimizuoti užduočių atlikimo laiką ir padidinti grįžtamojo ryšio kokybę;
- studentai gali savarankiškai įsitikinti teorinių teiginių teisingumu, tai skatina jų pažinimą, sužadina jų norą tirti ir atrasti mokslines tiesas;
- ugdomi studentų modeliavimo gebėjimai, pagilėja studentų statistinių metodų suvokimas.

Literatūra

1. J. Adomavičius ir kt. *Informatika 2: uždavinių paruošimas sprendimui Matlab ir Mathcad aplinkose*. Technologija, Kaunas, 2008.
2. P. Jucevičienė, V. Petkūnas. IKT diegimo įtakos pedagoginei sistemai kaitos charakteristikos. *Socialiniai mokslai*, 2(39):38–47, 2003.

REZIUMĖ

V.P. Pekarskas. Application of computers for teaching the design of experiments

A concrete example of application of computer technologies for teaching the course "Planning of Experiments" is analyzed in the paper. It is shown that the application of computer technologies does not only enrich the traditional educational process, but increases its effectiveness and widens teaching and learning possibilities.

Keywords: design of experiments, application of Mathcad.