

Statistično planiranje in vrednotenje metalurških raziskav

LATINSKI KVADRAT S PROGRAMOM NA RAČUNALNIKU ZUSE- Z 23

Latinski kvadrat je statistična metoda s področja analize variance. V članku je podrobno opisano, kako to metodo uporabljamo od zbiranja podatkov pa do tolmačenja rezultatov pri uporabi elektronskega računalnika. Značilno za latinski kvadrat je, da potrebujemo majhno število zbranih podatkov, čeprav opazujemo kar tri vplive hkrati. Zato pa morajo biti zbrani podatki čisto na poseben način v skladu z določenim modelom. Praktičen primer je vzet iz raziskovalnega arhiva železarne Ravne. Programa pa sta bila izdelana za računalnik Zuse Z-23 in računalnik IBM system/ 360 model 30.

UVOD

Latinski kvadrat je poseben primer analize variance. Z njim primerjamo hkrati različna stanja treh vplivnih faktorjev. Prednost latinskega kvadrata je v tem, da nam je treba zbrati le malo število podatkov. Ni potrebno zbrati podatkov za vse možne kombinacije stanj treh vplivnih faktorjev. Zberemo le toliko podatkov, kolikor je kombinacij stanj dveh vplivnih faktorjev. Seveda pa mora biti zato zbiranje podatkov točno vnaprej določeno. Vsi trije vplivni faktorji pa morajo imeti enako število stanj. Recimo, da je stanj vsakega vplivnega faktorja m (m je poljubno naravno število), potem bi bilo vseh kombinacij stanj treh vplivnih faktorjev $m \times m \times m$ (ali m^3). V resnici pa potrebujemo za latinski kvadrat le podatke za $m \times m$ (ali m^2) kombinacij stanj. Če je na primer $m = 8$, potrebujemo namesto 512 le 64 podatkov. To pa je velika razlika v stroških in času preiskave.

V tem članku si bomo ogledali zbiranje podatkov in tolmačenje rezultatov latinskega kvadrata za program, ki smo ga izdelali na računalniku ZUSE Z-23. Program za latinski kvadrat je bil izdelan tudi za računalnik IBM/360 v jeziku Fortran. Rezultati tega programa so popolnoma isti kot po programu za računalnik Z-23. Zato tolmačenje rezultatov, ki ga dobimo v tem članku, lahko uporabimo tudi za rezultate iz računalnika IBM/360. Priprava podatkov za računalnik IBM/360 pa se seveda razlikuje, ker uporabimo tam luknjane kartice, pri Z-23 pa teleprinterški trak. Program

smo uvedli pri raziskavah na področju luženja jekla pri ugotavljanju vplivov luženja na stanje površine. Zato se v nadaljevanju pri razlagi programa poslužujemo praktičnih podatkov te raziskave luženja.

PLAN ZBIRANJA PODATKOV

Za primer si vzemimo luženje določene vrste jekla. Predpostavimo, da vpliva na izgled površine po luženju čas in temperatura luženja ter vrsta lužila. To so trije vplivni faktorji. Takoj je dobro, da se domenimo za vrstni red vplivnih faktorjev. Prvi naj bo čas, drugi temperatura in tretji vrsta lužila. Ker imamo na razpolago osem vrst lužila, pravimo, da ima tretji vplivni faktor 8 stanj. Določeno stanje tretjega vplivnega faktorja pomeni uporabo določene vrste lužila. Ker mora biti pri latinskem kvadratu število stanj vseh treh vplivnih faktorjev enako, si moramo izbrati 8 časov in 8 temperatur, pri katerih bomo lužili. Za čase si izberimo:

- 1' 10 min.
- 2' 20 min.
- 3' 30 min.
- 4' 40 min.
- 5' 60 min.
- 6' 1 uro 15 min.
- 7' 1 uro 30 min.
- 8' 2 uri

Za temperature si izberimo:

- 1' 10° C
- 2' 20° C
- 3' 30° C
- 4' 40° C
- 5' 50° C
- 6' 60° C
- 7' 70° C
- 8' 80° C

Napišimo zaradi popolnosti še stanja tretjega vplivnega faktorja, to je sestava lužil, ki jih uporabimo pri poskusih:

- 1' 225 ml HCL conc.
25 ml HNO₃ conc.
250 ml vode

- 2' 90 ml H_2SO_4
25 ml nasičene $NaNO_3$
385 ml vode
- 3' 30 ml H_2SO_4
30 ml HF
90 ml HCL
350 ml vode
- 4' 25 ml H_2SO_4 conc.
50 ml HNO_3
50 ml HF
375 ml vode
- 5' 20 ml HNO_3
5 ml H_3PO_4
25 ml CH_3COOH
5 ml HCL
425 ml vode
- 6' 100 ml H_3PO_4
15 ml HNO_3
15 ml HCL
350 ml vode
- 7' 33 g $FeSO_4$
10 g HF
500 g vode
- 8' 50 g HNO_3
2,5 g HF
450 g vode

Podatke za izgled površine zberemo tako, da naredimo 64 prob, ki jih po poizkusu luženja med seboj primerjamo in določimo range, to je števila, ki nam po vrstnem redu ocene merijo izgled. To seveda nima zveze z metodo latinskega kvadrata. Tako dobimo le številčne podatke, ki jih potem med seboj primerjamo po metodi latinskega kvadrata.

8	5	6	1	4	3	7	2	①
7	4	5	6	3	2	1	8	②
3	7	1	2	8	4	6	5	③
1	2	7	4	6	5	8	3	④
2	8	3	5	1	6	4	7	⑤
4	6	8	3	5	7	2	1	⑥
5	1	4	7	2	8	3	6	⑦
6	3	2	8	7	1	5	4	⑧
①	②	③	④	⑤	⑥	⑦	⑧	

Stanje 2. vplivnega faktorja

Stanje 1. vplivnega faktorja

Slika 1

Primer latinskega kvadrata za določevanje vplivov luženja

Kako izberemo kombinacije treh vplivnih faktorjev za 64 prob? Naredimo shemo, ki je tipična za latinski kvadrat. (slika 1)

Vsak od 64 kvadratov v shemi na sliki 1 pripada eni kombinaciji treh vplivnih faktorjev: časa, temperature luženja in uporabi lužila. Na koncu vrste, v kateri opazovani kvadrataček leži, nam obkroženo število pove, za katero stanje 1. vplivnega faktorja gre. Na koncu stolpca, v katerem opazovani kvadrataček leži, nam obkroženo število pove, za katero stanje 2. vplivnega faktorja gre. V opazovanem kvadratku samem pa je napisano stanje 3. vplivnega faktorja. Vzemimo za primer v tretji vrsti sedmi kvadrataček. Ker leži v tretji vrsti, gre za stanje 3 1. vplivnega faktorja. To je, kakor vidimo iz prejšnjega seznama stanj, čas 30 minut. Ker leži kvadrataček v sedmem stolpcu, gre za stanje 7 2. vplivnega faktorja. To je temperatura $70^{\circ}C$. Ker pa v opazovanem kvadratku piše 6, to pomeni, da gre za šesto stanje 3. vplivnega faktorja, ki je v našem praktičnem primeru lužilo s sestavo: 100 ml H_3PO_4 , 15 ml HNO_3 , 15 ml HCL in 350 ml vode.

Vedno se moramo držati pravila, da se z vrsticami menjajo stanja 1. vplivnega faktorja, s stolpci pa stanja 2. vplivnega faktorja, ker je program za računalnik tako narejen.

Iz sheme na sliki 1 vidimo, da so števila, ki določajo stanja 3. vplivnega faktorja, v kvadratih zapisana tako, da niti v isti vrsti niti v isti koloni ne dobimo nikjer dveh enakih števil. To je princip sheme latinskega kvadrata. Ko zremo za število stanj vplivnih faktorjev, ki ga splošno označimo z m (glej uvod), narišemo najprej $m \times m$ kvadratov in vanje vpišemo števila od 1 do m , tako da nikoli ni v isti vrsti ali koloni dveh enakih števil. V začetku vpisovanja števil v kvadratke imamo precej svobode, nato pa moramo vedno bolj paziti, da se držimo omenjenega pravila. Vpisovanje števil v kvadratke naj bo čisto slučajno, ne da bi se pri tem držali kakega vrstnega reda. Obstajajo posebna pravila, kako shemo latinskega kvadrata čisto slučajnostno izberemo, vendar le do $m = 6$. Pri pisanju sheme ne smemo nikoli misliti na to, kako bomo določili probo za izbrana stanja vplivnih faktorjev, potem preizkusili in ali bo rezultat ugoden. Zato je najbolje, da shemo izberemo čisto neodvisno od narave poizkusa. Kot podatek je le število m . Vzemimo npr., da je $m = 5$. Nekdo bi čisto enostavno narisal 25 kvadratov in vanje napisal števila od 1 do 5 kakor kaže slika 2.

Sele potem bi vsaki vrstici pripisali eno stanje prvega vplivnega faktorja, vsaki koloni eno stanje drugega vplivnega faktorja in vsakemu številu od 1 do 5 eno stanje tretjega vplivnega faktorja. Vrnimo se na naš primer luženja. Po shemi na sliki 1 preizkusimo vseh 64 prob. Rezultate za izgled površine, iz vrednotene s števili od 1 do 64 (rangji), vpišemo v kvadratke, ki jim ustrezajo. Če ima več prob enako površino, dodelimo vsem enako število

4	3	2	5	1
1	5	4	2	3
5	4	3	1	2
3	2	1	4	5
2	1	5	3	4

Slika 2
Latinski kvadrat 5 × 5

ranga — to je sredina zasedenega območja rangov. Npr.: Če tri probe zasedajo z enako oceno mesta od 3—5, imajo vse tri probe rang 4, če pa zasedajo mesta od 3—6, imajo rang 4,5.

Tako dobimo shemo za latinski kvadrat na sliki 3.

		Temperatura [°C]							
		10	20	30	40	50	60	70	80
Čas [min.]	10	8 2,8	5 10,5	6 10,5	1 20	4 36,5	3 10,5	7 10,5	2 4,3
	20	7 10,5	4 51	5 20	6 20	3 20	2 43	1 6,2	8 36,5
	30	3 10,5	7 3,5	1 2,8	2 3,5	8 35,5	4 4,3	6 2,8	5 60
	40	1 20	2 36,5	7 10,5	4 51	6 20	5 51	8 36,5	3 60
	60	2 32	8 57	3 20	5 36,5	1 32	6 43	4 51	7 2,8
	1 ^h 15	4 57	6 20	8 51	3 20	5 51	7 3,5	2 4,3	1 6,4
	1 ^h 30	5 20	1 2,8	4 57	7 10,5	2 3,5	8 4,3	3 4,3	6 51
	2 ^h	6 3,5	3 3,5	2 51	8 32	7 20	1 6,3	5 60	4 51

Slika 3
Male številke pomenijo oceno izgleda površine v rangih.

Iz izpolnjene sheme latinskega kvadrata prepisemo podatke na teleprinter. Protokol podatkov vidimo na sliki 4.

8'
8' 28 5' 10,5 6' 10,5 1' 20 4' 36,5 3' 10,5 7' 10,5 2' 43
7' 10,5 4' 51 5' 20 6' 20 3' 20 2' 43 1' 62 8' 36,5
3' 10,5 7' 3,5 1' 28 2' 3,5 8' 36,5 4' 43 6' 28 5' 60
1' 20 2' 36,5 7' 10,5 4' 51 6' 20 5' 51 8' 36,5 3' 60
2' 32 8' 57 3' 20 5' 36,5 1' 32 6' 43 4' 51 7' 28
4' 57 6' 20 8' 51 3' 20 5' 51 7' 3,5 2' 43 1' 64
5' 20 1' 28 4' 57 7' 10,5 2' 3,5 8' 43 3' 43 6' 51
6' 3,5 3' 3,5 2' 51,0 8' 32 7' 20 1' 63 5' 60 4' 51
3' 10 1,865
5 2,24
1 3,1

Slika 4
Protokol podatkov za latinski kvadrat

Najprej samo stoji število 8', ki pomeni, da gre za latinski kvadrat 8 × 8. Nato je točno prepisana izpolnjena shema s slike 3, in sicer tako, da najprej napišemo za vsak kvadratega število, ki pomeni stanje tretjega vplivnega faktorja (velike številke na sliki 3) in nato šele vrednost podatka (male številke na sliki 3). Najprej napišemo prvo vrsto, nato drugo itd. Ko smo celo izpolnjeno shemo s slike 3 prepisali, slede še podatki za F test!

Najprej število 3', ki pomeni, da bomo testirali pomembnost razlik na treh nivojih $\alpha = 10$, $\alpha = 5$ in $\alpha = 1$. Števila 10,5 in 1 so napisana pred ustreznimi vrednostmi iz tabel za F_{α} : v_1 , v_2 . Prostostni stopnji v_1 in v_2 izračunamo po obrazcih

$$v_1 = m - 1$$

$$v_2 = (m-1) (m-2)$$

V našem primeru, ko je $m = 8$, torej poiščemo vrednosti iz tabel!

$$F_{10}; 7,42 = 1,865$$

$$F_5; 7,42 = 2,24$$

$$F_1; 7,42 = 3,10$$

in jih napišemo v zadnjem delu podatkov, kot to vidimo v primeru na sliki 4.

RAZLAGA REZULTATOV

Primer rezultatov po programu za latinski kvadrat vidimo na sliki 5.

Najprej je naslov »LATINSKI KVADRAT«, ki pomeni začetek rezultatov za en primer podatkov. Nato v treh vrstah sledijo srednje vrednosti. V prvi vrsti so srednje vrednosti izračunane za posamezne vrstice v shemi na sliki 3 — to je za posamezna stanja prvega vplivnega faktorja (časa). V drugi vrsti so srednje vrednosti, izračunane po stolpcih sheme na sliki 3 — to je za posamezna stanja drugega vplivnega faktorja (temperatura). V tretji vrsti rezultatov so srednje vrednosti za stanja tretjega vplivnega faktorja (lužila), in sicer po vrsti, kot smo lužila označili s števili od 1' do 8'.

Računalnik vedno vzame za prvi vplivni faktor tistega, ki se spreminja od vrstice do vrstice v shemi latinskega kvadrata; za drugi vplivni faktor pa tistega, ki se spreminja od stolpca do stolpca v shemi. Prvi vplivni faktor v rezultatih označi z 1', drugi vplivni faktor označi z 2' in tretji vplivni faktor s 3'.

LATINSKI KVADRAT

21.1875 32.8750 26.6250 35.6875 37.4375 38.6875 32.0000 35.5000
 22.6875 26.2500 31.0000 24.1875 27.4375 37.5000 41.7500 49.1875
 39.6250 31.9375 23.4375 49.6875 38.6250 24.5000 12.1250 40.0625

1957.6250 7 279.6607 1.78
 4970.8750 7 710.1250 4.52
 8019.7500 7 1145.6786 7.29
 6600.7500 42 157.1607

21549.0000 63

ALFA = 10.0

2'

8' 1' 26.5000

8' 2' 22.9375

8' 4' 25.0000

3'

1' 7' 27.5000

4' 3' 26.2500

4' 6' 25.1875

4' 7' 37.5625

5' 7' 26.5000

8' 7' 27.9375

ALFA = 5.0

2'

8' 1' 26.5000

8' 4' 25.0000

3'

1' 7' 27.5000

4' 3' 26.2500

4' 6' 25.1875

4' 7' 37.5625

5' 7' 26.5000

8' 7' 27.9375

ALFA = 1.0

2'

3'

4' 7' 37.5625

Slika 5

Protokol rezultatov računalnika

Za srednjimi vrednostmi v rezultatih sledi standardna shema za analizo variance z izračunanimi vrednostmi F v zadnji koloni. Cela števila v drugi koloni so ustrezne prostostne stopnje. Vsak F ima

enako število prostostnih stopenj ($\nu_1 = 7$ in $\nu_2 = 42$), ker so vsi trije vplivni faktorji enakovredni po številu zbranih podatkov. Pomembne razlike med srednjimi vrednostmi stanj posameznega vplivnega faktorja dobimo kot pri analizi variance za vsak nivo posebej. Najprej računalnik izpiše razlike, ki so 90 % pomembne torej $\alpha = 10$. V našem primeru izpiše za tekstom »ALFA = 10.0« najprej število 2', ki pomeni, da slede razlike med stanji drugega vplivnega faktorja — to je v našem primeru temperature. Nato sledi izpis v treh kolonah. V prvi koloni je vedno število, ki označuje stanje, katerega srednja vrednost je pomembno večja od srednje vrednosti stanja, ki ga označuje število v drugi koloni. V tretji koloni je razlika teh srednjih vrednosti.

V našem primeru ima stanje 8' 90 % pomembno večjo srednjo vrednost od stanj 1', 2' in 4'. To pomeni, da pri temperaturi 80° C dobimo drug izgled kot pri temperaturah 10° C, 20° C in 40° C. Za razlikami stanj drugega vplivnega faktorja slede razlike stanj tretjega vplivnega faktorja, ki ga označi število 3'. Razlaga razlik je enaka kot pri drugem vplivnem faktorju. Pomembnih razlik prvega vplivnega faktorja ni.

Za razlikami, ki so 90 % pomembne, slede 95 % pomembne razlike pod tekstom »ALFA = 5.0«. Vidimo, da je odpadla ena od razlik za drugi vplivni faktor.

Pod »ALFA = 1.0« slede še 99 % pomembne razlike. Tu vidimo, da za drugi vplivni faktor dobimo le izpis 2' in nič drugega. To pomeni da 99 % pomembnih razlik ni med dvema stanjema temperature. So pa tako imenovani kontrasti ali »razlike« med več kot dvema stanjema¹.

Za tretji vplivni faktor je ostala le še razlika med stanji 4' in 7'. Torej lahko trdimo z 99 % gotovostjo, da se sestavi lužil 4' in 7' razlikujeta po vplivu na izgled površine.

Literatura

1. Ing. B. Rode, Ing. J. Rodič, Statistično planiranje in vrednotenje metalurških raziskav, *Železarski zbornik*; II./1968 št. 2, str. 99—111

ZUSAMMENFASSUNG

Das lateinische Quadrat ist eine statistische Methode der Varianzen Analyse. Im Artikel ist die Anwendung dieser Methode, von der Sammlung der Daten, bis zu der Erläuterung der Ergebnisse, beim Gebrauch einer Elektronenrechenmaschine, ausführlich beschrieben. Charakteristisch für das lateinische Quadrat ist, dass wir eine nur geringe Zahl der Daten gebrauchen, wenn wir auch drei verschiedene Einflüsse zugleich beobachten.

Die Gesammelten Daten müssen aber auf eine ganz besondere Art mit dem bestimmten Modell übereinstimmen. Dieser praktische Beispiel ist aus dem Archiv der Versuchsabteilung des Hüttenwerkes Ravne entnommen worden. Die beiden Programme wurden mittels der Elektronenrechenmaschinen Zuse Z-23 und IBM/360 Modell 30 ausgefertigt.

SUMMARY

Latin square is a statistical method from the field of analysis of variance. In the paper the use of this method, is detailedly described from collecting of data to the interpretation of results using a computer. Characteristic of the latin square is that a small number of collected data is necessary

though simultaneously three trends are observed. Therefore the data must be collected in a special way, corresponding to a certain model. An example is taken from experimental records of Ravne Ironworks. The programs were made for the computer Zuse Z-23 and the computer IBM system 360 model 30.

ЗАКЛЮЧЕНИЕ

Латинский квадрат это статистический метод из области анализа регрессии. В статье подробно описано как употребить этот метод начиная от сбора данных до объяснения результатов при употреблении электронного счётчика. Характерно для латинского квадрата то, что для него достаточно небольшое количество

данных, несмотря на то, что в тоже самое время рассматриваются три влияния. Поэтому сбор сведений надо выполнить определенным способом в согласии специального модели. Из промышленной практики взят пример из архива металлургического завода Равне. Программы были приготовлены для счётчика Zuse Z-25 и для счётчика система IBM, 360, модель 30.