UNIVERZITA KOMENSKÉHO V BRATISLAVE
FAKULTA MATEMATIKY, FYZIKY A INFORMATIKY

Riadkovo-stĺpcové návrhy štatistických experimentov

BAKALÁRSKA PRÁCA

2020 Róbert Druska

UNIVERZITA KOMENSKÉHO V BRATISLAVE
FAKULTA MATEMATIKY, FYZIKY A INFORMATIKY

Riadkovo-stĺpcové návrhy štatistických experimentov

BAKALÁRSKA PRÁCA

Študijný program:	Matematika
Študijný odbor:	1114 Matematika
Školiace pracovisko:	Katedra aplikovanej matematiky a štatistiky
Vedúci práce:	doc. Mgr. Radoslav Harman, PhD.

Bratislava 2020

Róbert Druska

34792200

Univerzita Komenského v Bratislave
Fakulta matematiky, fyziky a informatiky

ZADANIE ZÁVEREČNEJ PRÁCE

Meno a priezvisko študenta: Róbert Druska
Študijný program: matematika (Jednoodborové štúdium, bakalársky I. st., denná forma)
Študijný odbor: matematika
Typ záverečnej práce: bakalárska
Jazyk záverečnej práce: slovenský
Sekundárny jazyk: anglický

Názov: Riadkovo-stĺpcové návrhy štatistických experimentov
Row-column designs of statistical experiments

Anotácia: Prvým cieľom je analyzovať vlastnosti odhadov parameterov regresného modelu pre takzvaný riadkovo-stĺpcový experimentálny návrh. Druhým cieľom je navrhnúť algoritmus na výpočet optimálneho návrhov tohto typu, v závislosti na požiadavkách experimentátora.

Vedúci: doc. Mgr. Radoslav Harman, PhD.
Katedra: FMFI.KAMŠ - Katedra aplikovanej matematiky a štatistiky
Vedúci katedry: prof. RNDr. Marek Fila, DrSc.

Dátum zadania: 15.10.2019

Dátum schválenia: 18.10.2019

prof. RNDr. Ján Filo, CSc.
garant študijného programu

....

študent vedúci práce

Poďakovanie TODO

Abstrakt

TODO

Kľúčové slová: lineárny regresný model

Abstract

TODO

Keywords: linear regression

Obsah

Úvod

TODO

1 Maticová príprava

V našej práci budeme často pracovať s poznatkami z lineárnej algebry, preto v krátkosti zhrnieme niektoré najdôležitejšie z nich.

Definícia 1.1. *Nech A je ľubovoľná matica typu $m \times n$. Potom symbolom $\mathcal{M}(A)$ označujeme stĺpcový priestor matice A, t. j.*

$$\mathcal{M}(A) = \{Au : u \in \mathbb{R}^n\}$$

Symbolom $\mathcal{K}(A)$ označujeme jadro matice (kernel), t. j.

$$\mathcal{K}(A) = \{v \in \mathbb{R}^n : Av = 0\}.$$

Lema 1.2. *Platí*

$$\mathcal{K}(A) = [\mathcal{M}(A)]^{\perp} \equiv \{v \in \mathbb{R}^n : \forall_{u \in \mathcal{M}(A)} u^T v = 0\}$$

Dôkaz: $x \in \mathcal{K}(A^T) \Leftrightarrow A^T x = 0 \Leftrightarrow \forall_v v^T A^T x = 0 \Leftrightarrow \forall_v (Av)^T x = 0 \Leftrightarrow x \in [\mathcal{M}(A)]^{\perp}$

Lema 1.3. *Nech A je ľubovoľná matica typu $m \times n$. Potom*

$$\mathcal{M}(A^T) = \mathcal{M}(A^T A)$$

Dôkaz: Nech $z \in \mathcal{M}(A^T A)$. Potom existuje také v, že $z = A^T A v = A^T u$ pre $u = Av$. Preto $z \in \mathcal{M}(A^T)$ a teda $\mathcal{M}(A^T A) \subseteq \mathcal{M}(A^T)$.

Teraz nech $z \in \mathcal{K}(A^T A)$. Potom:

$$A^T A z = (0, 0, \ldots, 0)^T \Rightarrow z^T A^T A z = 0 = (Az)^T A z = ||Az|| \Rightarrow Az = 0$$

Preto $z \in \mathcal{K}(A)$, a teda $\mathcal{K}(A^T A) \subseteq \mathcal{K}(A)$. Z lineárnej algebry ale vieme, že nulový priestor matice je kolmý doplnok jej riadkového priestoru, preto predošlý výraz vieme rozšíriť:

$$[\mathcal{M}(A^T A)]^{\perp} = \mathcal{K}(A^T A) \subseteq \mathcal{K}(A) = [\mathcal{M}(A^T)]^{\perp}$$

Odstránením $\perp$ sa inklúzia obráti, čo spolu s opačnou inklúziou vyššie dáva výsledok $\mathcal{M}(A^T) = \mathcal{M}(A^T A)$.

Definícia 1.4. *Nech A je ľubovoľná matica. Potom matica A^- taká, že*

$$AA^-A = A$$

sa nazýva g-inverzia (alebo pseudoinverzia) matice A.

Poznámka: Ak A^{-1} neexistuje, tak pre maticu A existuje nekonečný počet g-inverzií.

Pseudoinverzie použijeme neskôr v našej práci, keď budeme skúmať, či vektor patrí do stĺpcového priestoru matice. Na to nám poslúži nasledovná veta.

Veta 1.5. *Nech rovnica $Ax = y$ má riešenie a nech A^- je ľubovoľná g-inverzia. Potom A^-y je riešením tejto rovnice.*

Dôkaz: Keďže $Ax = y$ má riešenie, tak existuje také x_0, že $Ax_0 = y$. Potom

$$A(A^-y) = A(A^-Ax_0) = Ax_0 = y$$

1.1 Projekčné matice

Pri práci s lineárnym regresným modelom sa nám zídu poznatky o projekčných maticiach, napr. pri metóde najmenších štvorcov.

Definícia 1.6. *Nech $\mathcal{V} \subset \mathbb{R}^n$ je lineárny priestor. Potom matica P je lineárny projektor na $\mathcal{V}$ práve vtedy, keď platia nasledovné podmienky:*

a*)* $\forall_{x\in\mathbb{R}^n} Px \in \mathcal{V}$

b*)* $\forall_{v\in\mathcal{V}} Pv = v$

Z vlastností a) a b) vyplýva, že $P^2 = P$, t. j. matica P je idempotentná. Poznamenávame, že aj naopak, ak matica P je idempotentná a definujeme $V = \mathcal{M}(P)$, tak P má vlastnosti a) a b) z predošlej definície.

Takto definovaný projektor nemusí byť ortogonálny, t. j. nezobrazuje vektory na k nim najbližší vektor z priestoru V. Formálne ortogonalitu definujeme nasledovne.

Definícia 1.7. *Projektor P na lineárny podpriestor V je ortogonálny vzhľadom na skalárny súčin $\langle\cdot,\cdot\rangle$, ak $\langle x, Py\rangle = \langle Px, y\rangle$ pre všetky $x, y \in \mathbb{R}^n$.*

Veta 1.8. *Nech $F \in \mathbb{R}^{m\times m}$, $m > n$ má plnú hodnosť a nech $\rangle a, b\langle = a^Tb$ je skalárny súčin v $\mathbb{R}^m$. Potom ortogonálny projektor na $\mathcal{M}(F)$ je*

$$P = F(F^TF)^{-1}F^T$$

Dôkaz: Ukážeme, že P spĺňa vlastnosti z definície lineárneho projektora, aj podmienku ortogonality z definície 1.7.

$$\forall_{x\in\mathbb{R}^m} Px = F(F^TF)^{-1}F^Tx = Fu \in \mathcal{M}(F)$$

$$\forall_{l=Fu\in\mathcal{M}(F)} Pl = P(Fu) = F(F^TF)^{-1}F^TFu = Fu = l$$

$$\forall_{x,y\in\mathbb{R}^m} \langle x, Py\rangle = x^TF(F^TF)^{-1}F^Ty = x^T[F(F^TF)^{-1}F^T]^Ty = x^TP^Ty = \langle Px, y\rangle$$

Veta 1.9. *Predchádzajúca veta platí aj pre maticu F, ktorá nemá plnú hodnosť, ak namiesto inverzie $(F^TF)^{-1}$ použijeme ľubovoľnú g-inverziu $(F^TF)^-$.*

2 Lineárny regresný model

Lineárna regresia patrí v súčasnosti medzi najpoužívanejšie štatistické metódy. Aplikácie regresie možno nájsť v mnohých vedeckých odvetviach, vrátane medicíny, biológie, ekonómie, sociológie atď. Ciele regresnej analýzy zväčša možno zaradiť do troch oblastí:

1. nájsť vzťah medzi závislou premennou y a regresormi $x_1, \ldots, x_n$,
2. odhadnúť y na základe hodnôt $x_1, \ldots, x_n$,
3. preskúmať premenné $x_1, \ldots, x_n$ a identifikovať, ktoré z nich lepšie popisujú premennú y, resp. určiť, ktoré z nich majú na y aký vplyv

V našej práci sa budeme zaoberať treťou z pomenovaných oblastí.

Uvažujme teda lineáry regresný model

$$y = Xb + \varepsilon \tag{1}$$

Na ľavej strane y je $n \times 1$ vektor pozorovaných náhodných premenných. Vektor y nazývame *závislou* premennou. Na pravej strane X je matica plánu, b je náhodný vektor lineárneho vzťahu a ε je náhodná chyba. X vo všeobecnosti môže, ale nemusí byť náhodná premenná. Pre účely našej práce budeme narábať s fixnou (známou) maticou X, čo zohľadníme aj v ďalších teoretických základoch v tejto kapitole.

Nech teda X je známa matica $n \times p$ tvaru:

$$X = \begin{bmatrix} 1 & x_{11} & \ldots & x_{1p} \\ \vdots & \vdots & \ddots & \\ 1 & x_{n1} & \ldots & x_{np} \end{bmatrix}$$

Rovnica (1) nám určuje n lineárnych vzťahov tvaru:

$$y_i = b_0 + b_1 x_{i1} + \ldots + b_p x_{ip} + \varepsilon_i = b^T x_i + \varepsilon_i$$

pre $y = (y_1, y_2, \ldots, y_n)^T$, $b = (b_1, b_2, \ldots, b_p)^T$ a $\varepsilon = (\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n)^T$.

Ďalej budeme predpokladať $E(\varepsilon) = \mathbf{0}$, $Var[\varepsilon] = \sigma^2 I$, t. j jednotlivé $\varepsilon_i, \varepsilon_j$ sú nezávislé a rovnako rozdelené. Pre jednoduchosť takisto predpokladajme, že $\sigma^2 = 1$.

V praxi je b neznámy vektor, ktorý sa snažíme odhadnúť na základe nameraných hodnôt y. Na spočítanie odhadu sa používajú rôzne metódy, najčastejšie napr. metóda najmenších štvorcov alebo metóda maximálnej vierohodnosti.

2.1 Metóda najmenších štvorcov

Metódou najmenších štvorcov vypočítame odhad $\hat{b}$ parametra b nasledovne:

$$\hat{b} = \underset{b \in \mathbb{R}^p}{\operatorname{argmin}} (y - Xb)^T C (y - Xb) = \underset{b \in \mathbb{R}^p}{\operatorname{argmin}} ||y - Xb||^2_{C^{-1}}$$

kde C je nejaká kladne definitná matica. Ak $C = I$, potom minimalizujeme výraz $||y - Xb||^2_I = ||y - Xb||^2 = \sum_{i=1}^n (y_i - X_{i.}b)^2$, kde $X_{i.}$ značí i-ty riadok matice X. Keďže sme predpokladali nezávislosť chýb a tiež homogenitu ich rozptylu, skutočne je pre nás najvýhodnejšie dosadiť $C = I$. Preto maticu C v ďalšom opise teórie spomínať nebudeme.

Geometricky metódu najmenších štvorcov možno interpretovať ako projekciu vektora y na stĺpcový priestor matice X. Hľadáme teda taký vektor $\hat{b}$, pre ktorý platí $X\hat{b} = Py$, kde P je matica ortogonálnej projekcie na stĺpcový priestor X. Z vety (1.9) vieme, že $P = X(X^T X)^- X^T$.

Odhad $\hat{b}$ parametra b je teda riešením rovnice

$$Xb = X(X^T X)^- X^T y \tag{2}$$

kde $P = X(X^T X)^- X^T$ je ortogonálny projektor na stĺpcový priestor X.

Všetky riešenia tejto rovnice musia mať tvar:

$$(X^T X)^- X^T y$$

kde použitá g-inverzia je ľubovoľná.

Z uvedeného vyplýva, že v prípade regulárnosti $X^T X$ je odhad $\hat{b}$ parametra b jednoznačný. V našej práci budeme skúmať matice (modely) X, ktoré nie sú regulárne, takže jednoznačný odhad $\hat{b}$ nebudeme schopní nájsť (čo v konečnom dôsledku ani nie je naším záujmom). Budeme odhadovať lineárnu funkciu zložiek vektora b, konkrétne $h^T b = h_1 b_1 + \ldots + h_p b_p$, ktorá býva jednoznačne odhadnuteľná aj v prípade singularity $X^T X$, ak vektor h spĺňa určité podmienky.

Definícia 2.1. *Lineárnu kombináciu $h^T b$ zložiek vektora b nazývame odhadnuteľnou, ak pre ľubovoľné riešenia b^* a b^{**} rovnice (2) platí $h^T b^* = h^T b^{**}$.*

Je niekoľko ekvivalentných podmienok, ktoré stačia na to, aby $h^T b$ bolo odhadnuteľné. Z nich spomenieme dve v nasledovnej vete, ktorú použijeme neskôr v našej práci.

Veta 2.2. *$h^T b$ je odhadnuteľné, ak platia nasledovné ekvivalentné podmienky:*

1. $h \in \mathcal{M}(X^T)$

2. $h \in \mathcal{M}(X^T X)$,

kde $\mathcal{M}$ označuje stĺpcový priestor matice.

Dôkaz: Ukážeme, že ak $h \in \mathcal{M}(X^T)$ tak $h^T b$ je odhadnuteľné. Nech teda $h \in \mathcal{M}(X^T)$. Potom existuje také u, že $h = X^T u$. Preto pre ľubovoľné riešenie $\hat{b}$ rovnice (2) platí:

$$h^T \hat{b} = u^T X \hat{b} = u^T P y$$

kde päta kolmice Py je jednoznačne daná, preto aj $h^T \hat{b}$ je jednoznačne dané (bez ohľadu na voľbu $\hat{b}$). $h^T b$ je teda podľa definície (2.1) odhadnuteľné.

Ekvivalencia podmienok (1) a (2) vyplýva z toho, že podľa lemy (1.3) vo všeobecnosti platí, že $\mathcal{M}(X^T) = \mathcal{M}(X^T X)$. $\square$

Ak h patrí do riadkového priestoru matice X, potom existuje také u, že $h = X^T u$. Potom pre jednoznačný odhad $h^T \hat{b}$ vektora $h^T b$ platí:

$$h^T \hat{b} = u^T X \hat{b} = u^T P y = u^T X (X^T X)^- X^T y$$

Vidíme, že v strede výrazu na pravej strane je projekčná matica $X(X^T X)^- X^T$, ktorá je vždy jednoznačne určená, nezávisle na voľbe zovšeobecnenej g-inverzie v danom výraze.

Výsledok predchádzajúcej vety je dôležitý pre našu prácu, pretože nebudeme skúmať odhady b, ale odhady niektorých lineárnych kombinácií zložiek vektora b, konkrétne napr. rozdiely medzi parametrami.

2.2 Gaussova-Markovova veta

V predchádzajúcej časti sme ukázali, že ak matica je matica plánu X regulárna, jej odhad metódou najmenších štvorcov je rovný

$$\hat{b} = \underset{b \in \mathbb{R}^p}{\operatorname{argmin}} ||y - Xb||^2 = (X^T X)^{-1} X^T y$$

Pre strednú hodnotu odhadu $\hat{b}$ teda platí:

$$E[\hat{b}] = E[(X^TX)^{-1}X^Ty] = (X^TX)^{-1}X^TE[y] = (X^TX)^{-}X^TE[Xb+\varepsilon] =$$
$$(X^TX)^{-}X^TXb = b$$

To znamená, že $\hat{b}$ je lineárnym nevychýleným odhadom parametra b. Pre varianciu $\hat{b}$ dostávame:

$$Var[\hat{b}] = Var[(X^TX)^{-1}X^Ty] = (X^TX)^{-1}X^TVar[Xb+\varepsilon]X(X^TX)^{-1} =$$
$$(X^TX)^{-1}X^TVar[\varepsilon]X(X^TX)^{-1} = (X^TX)^{-1}X^TIX(X^TX)^{-1} = (X^TX)^{-1}.$$

Hodnota disperzie $(X^TX)^{-1}$ je v určitom zmysle najmenšia možná, t. j. $\hat{b}$ je najlepší lineárne nevychýlený odhad. Podobné tvrdenie platí pre odhad $h^T\hat{b}$ parametra h^Tb, ktorý je takisto lineárne nevychýlený. Upresníme to v nasledujúcej vete.

Neskôr v našej práci budeme hľadať také modely X, pri ktorých je disperzia odhadov h^Tb najmenšia možná, čo nám dá najlepší lineárny nevychýlený odhad. Ak nami navrhované modely X budú opisovať ten istý experiment, ten model X, pre ktorý disperzia odhadu h^Tb bude najmenšia, bude svojím spôsobom optimálny.

K nájdeniu optimálneho modelu X nám poslúži Gaussova-Markovova veta, ktorá určuje minimálnu možnú disperziu odhadu h^Tb.

Veta 2.3. *(Gaussova-Markovova) Nech h je z riadkového priestoru X. Potom minimálna možná disperzia lineárneho nevychýleného odhadu h^Tb je*

$$m = Var[h^T\hat{b}] = h^TM^-h$$

kde $M = X^TX$ je informačná matica parametra b a M^- je jej ľubovoľná g-inverzia a $\hat{b}$ je ľubovoľné riešenie rovnice

$$(X^TX)b = X^Ty$$

3 Metodika skúmania

Cieľom nášho skúmania bude nájsť ideálnu maticu vstupu do experimentu (konkrétny problém popíšeme v nasledujúcej kapitole). Pokúsime sa vysloviť závery pre všeobecný prípad matice $m \times n$, avšak pre istý prvotný náhľad do problematiky problém popíšeme a preskúmame na menších rozmeroch, konkrétne do rozmeru, do ktorého nám to umožní výpočtová technika. Pracovať budeme s programovacím jazykom *python*.

Od „preskúmania situácie“ do istého rozmeru si sľubujeme dôležité náhľady, ktoré nám umožnia vysloviť závery a hypotézy pre všeobecný prípad $m \times n$. Tie sa následne pokúsime dokázať matematickými metódami.

4 Riadkovo-stĺpcový aditívny model s dvoma typmi ošetrení

Cieľom tejto práce je skúmať experiment reprezentovaný binárnou maticou. Možná reprezentácia nášho modelu pri matici rozmeru $m \times n$ je takáto:

Uvažujeme model s dvomi kvalitatívnymi faktormi A a B, pričom faktor A má m úrovní a faktor B n úrovní. Pre každú kombináciu faktorov experimentátor zvolí jedno z dvojice ošetrení (angl. treatment). Budeme uvažovať len aditívny model bez interakcií.

Napríklad faktor A môže označovať dobrovoľníka (t.j. máme m dobrovoľníkov) a faktor B účinnú látku, ktorú nazveme liečivo (t.j. máme n liečiv). Ošetrenie 0 reprezentuje podanie liečiva orálnym spôsobom a ošetrenie 1 vnútrožilne). Účinok, ktorý po čase nameriame na dobrovoľníkovi, závisí aditívne od efektu samotného dobrovoľníka (jeho predispozícií), efektu liečiva a efektu ošetrenia. Všetky tieto efekty uvažujeme ako neznáme parametere modelu.

Návrh teda možno reprezentovať binárnou maticou (alebo bipartitným grafom, čo opíšeme neskoršie v našej práci).

Cieľom experimentu je odhadnúť rozdiel medzi dvoma efektmi (treatmentami). Cieľom hľadania optimálneho modelu je nájsť taký návrh, pri ktorom rozptyl Gauss-Markovovho odhadu rozdielu medzi efektmi dvojice ošetrení (t.j. odhadu kontrastu ošetrení) bude čo najmenší. (Použijeme pri tom Gaussovu-Markovovu vetu z predošlej kapitoly.)

Pre maticu typu $m \times n$ dostaneme teda mn vzťahov tvaru:

$$y_{ij} = a_i + b_j + t + e_{ij}$$

$i = 1, \ldots, m$; $j = 1, \ldots, n$; t je t_0 alebo t_1 podľa hodnoty na ij-tej pozícii matice a e_{ij} je chyba. Budeme odhadovať rozdiel medzi t_0 a t_1.

Z návrhu modelu v podobe binárnej matice tvaru $m \times n$ vytvoríme lineárny regresný model s maticou tvaru $mn \times (m+n+2)$, pretože máme mn meraní závislých od $m+n+2$ premenných.

Príklad:

Maticu návrhu experimentu

$$\begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \tag{3}$$

prepíšeme ako:

$$\begin{bmatrix} 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 1 & 1 & 0 \end{bmatrix} \tag{4}$$

Ako presne sme pri generovaní matice postupovali? Každý riadok reprezentuje jednu z hodnôt v matici (3), preto máme dokopy $3 \times 3 = 9$ riadkov. Prvé tri stĺpce reprezentujú riadok, v ktorom sa hodnota nachádza, ďalšie tri reprezentujú stĺpec. Vo všeocenosti ij-tu pozíciu matice typu $m \times n$ reprezentuje riadok, ktorý má jednotky na i-tej a $(m + j)$-tej pozícii. Posledné dva stĺpce sú určené tým, či je hodnota 0 alebo 1.

Označme vo všeobecnosti maticu modelu (3) B a maticu lineárneho zobrazenia (4) X. Predpokladáme lineárne vzťahy, preto dostávame lineárnu regresiu

$$y = Xb + e$$

kde y je vektor dát nameraných v experimente, b je vektor neznámych koeficientov a $e = (e_{11}, e_{12}, e_{13}, \ldots, e_{m1}, \ldots, e_{mn})^T$ je vektor chýb merania. Uvedomme si, že v našej práci nebudeme pracovať s konkrétnym vektorom y; snažíme sa totiž navrhnúť maticu modelu B (a z nej vyplývajúcu maticu X) ešte pred tým, než experiment vykonáme.

Naším cieľom teraz bude odhadnúť rozdiel medzi efektmi t_0 a t_1, resp. zistiť, aká môže byť disperzia tohto odhadu pri danej matici návrhu B. Modely, pre ktoré bude možná disperzia najmenšia, budeme považovať za optimálne.

V danej lineárnej regresii $y = Xb + e$ teda nebudeme odhadovať celý vektor b, ale len jeho lineárnu kombináciu $h^T b$, kde h je vektor rozmeru $m + n + 2$, ktorý má na $m + n$ miestach 0 a na zvyšných dvoch miestach $+1$ a -1, ktoré zodpovedajú efektom t_0 a t_1. Z algoritmu, ktorým sme z matice návrhu B vytvorili maticu lineárnej regresie X, vyplýva, že vektor h má tvar $(0, 0, 0, \ldots, +1, -1)^T$.

Na zistenie, či $h^T b$ je odhadnuteľné, použijeme vetu (2.2), a na následné nájdenie minimálneho rozptylu Gauss-Markovovho odhadu rozdielu medzi efektami použijeme Gaussovu-Markovovu vetu (2.3).

4.1 Odhadnuteľnosť $h^T b$

Pokúsime sa preskúmať, kedy matica návrhu umožňuje odhadnuteľnosť hodnoty $h^T b$ pre vyššie spomenuté h. Situáciu preskúmame do rozmeru 4×5 a na základe našich zistení vyslovíme hypotézu pre všeobecný prípad $m \times n$.

Postupovať budeme nasledovne: vyberieme si malý rozmer a pre všetky matice daného rozmeru zistíme, či $h^T b$ bude alebo nebude odhadnuteľné, a to nasledovným spôsobom. Z danej matice návrhu B vytvoríme maticu lineárnej regresie X spôsobom spomenutým vyššie. Hodnota $h^T b$ bude na základe vety (2.2) odhadnuteľná práve vtedy, keď vektor $h = (0, 0, 0, \ldots, +1, -1)^T$ patrí do stĺpcového priestoru matice $X^T X$.

Označme $M = X^T X$, túto maticu budeme nazývať informačnou maticou. Ak h patrí do stĺpcového priestoru M, potom existuje taký vektor v, že $Mv = h$. Na základe vety (1.5) vieme, že ak toto riešenie existuje, tak sa rovná $M^- h$, pričom zvolená g-inverzia môže byť ľubovoľná. Preto na zistenie, či h patrí do stĺpcového priestoru $M = X^T X$ stačí overiť platnosť rovnosti $M(M^- h) = h$.

Demonštrujme algoritmus na maticiach rozmeru 3×3. Existuje $2^9 = 512$ binárnych matíc typu 3×3, a keď sme na nich rozbehli daný algoritmus, dospeli sme k zisteniu, že pri 12 z nich hodnota $h^T b$ NIE JE odhadnuteľná. To znamená, že optimálne matice návrhu budeme hľadať v zvyšných 500 maticiach.

Ako vyzerajú matice, pri ktorých $h^T b$ nie je odhadnuteľné? Zoznam všetkých sa nachádza v Prílohe A, tu uvedieme 3 z nich:

$$\begin{bmatrix}1&1&1\\0&0&0\\1&1&1\end{bmatrix}, \begin{bmatrix}0&0&1\\0&0&1\\0&0&1\end{bmatrix}, \begin{bmatrix}0&0&0\\0&0&0\\0&0&0\end{bmatrix}$$

Týmto spôsobom sme pri našom výskume získali zoznam matíc, pri ktorých $h^T b$ nie je odhadnuteľné, až do rozmeru 4×5. Zoznam všetkých sa nachádza v GitHub repozitári spomenutom v úvode. Na základe podobnosti matíc daného typu sme prišli s nasledovnou hypotézou.

Tvrdenie 4.1. *$h^T b$ je odhadnuteľné práve vtedy, keď v matici návrhu B existuje aspoň jeden riadok a aspoň jeden stĺpec taký, ktorý má aspoň jednu 0 a aspoň jednu 1.*

Dôkaz: Na základe vety (2.2) vieme, že $h^T b$ je odhadnuteľné práve vtedy, keď $h \in \mathcal{M}(X^T)$ (keď h patrí do riadkového priestoru X), preto s danými podmienkami môžeme pracovať ekvivalentne.

$\boxed{\Longrightarrow}$ Nech $h = (0, \dots, +1, -1) \in \mathcal{M}(X^T)$. Použijeme dôkaz sporom. Nech v matici B typu $m \times n$ neexistuje taký riadok, ktorý má aspoň jednu 0 a aspoň jednu 1, t. j všetky riadky obsahujú buď samé 0 alebo samé 1. Ako vyzerá matica X?

Pozrime sa na riadky matice X zodpovedajúce prvému riadku matice B. Sú to práve tie riadky, ktoré majú na prvej pozícii 1, pričom si uvedomme, že všetky ostatné riadky majú na prvej pozícii 0.

$$x_1 = (1, 0, \dots, 1, 0, \dots, 0, 1, 0),$$

$$x_2 = (1, 0, \dots, 0, 1, \dots, 0, 1, 0),$$

$$\vdots$$

$$x_n = (1, 0, \dots, 0, 0, \dots, 1, 1, 0)$$

Keďže $h \in \mathcal{M}(X^T)$, existuje taká lineárna kombinácia riadkov X, ktorá dokopy dáva vektor h. Aké môžu mať v tejto lineárnej kombinácii zastúpenie riadky x_1, ..., x_n? Označme postupne a_1, ..., a_n koeficienty vektorov x_1, ..., x_n v lineárnej kombinácii riadkov X, ktorá dáva vektor h. Keďže x_1, ..., x_n sú jediné riadky matice X s 1 na prvej pozícii a vektor h má na prvej pozícii 0, musí platiť $\sum_{i=1}^{n} a_i = 0$.

Z predpokladu, že každý z riadkov matice B obsahuje buď samé 0 alebo samé 1, vyplýva, že každý z vektorov x_1, ..., x_n má rovnaké posledné dve hodnoty, a to buď $(0,1)$ alebo $(1,0)$. Z toho spolu s rovnosťou $\sum_{i=1}^{n} a_i = 0$ vyplýva, že lineárna kombinácia $\sum_{i=1}^{n} a_i x_i$ bude mať na posledných dvoch miestach hodnoty $(0,0)$, a teda nijako "neprispeje" do vektoru h. Vektor h teda nemožno napísať ako lineárnu kombináciu riadkov X, čo je podmienka odhadnuteľnosti.

Analogicky môžeme postupovať pre všetky riadky matice B. Týmto spôsobom pokryjeme všetky riadky matice X, vďaka čomu dospejeme k záveru, že z riadkov X nie je možné lineárnou kombináciou zostrojiť vektor h. To nám dáva spor s predpokladom $h \in \mathcal{M}(X^T)$. Preto nemôže platiť, že všetky riadky matice B obsahujú buď samé 0 alebo samé 1, a platí opačné tvdenie, t.j v B existuje aspoň jeden riadok taký, ktorý obsahuje 0 aj 1.

Rovnakou úvahou pre stĺpce B dospejeme k záveru, že B musí takisto obsahovať aspoň jeden stĺpec obsahujúci 0 aj 1.

$\boxed{\Longleftarrow}$ Nech matica B je taká, že existuje aspoň jeden riadok a aspoň jeden stĺpec také, že majú 0 aj 1. Zostrojíme takú lineárnu kombináciu riadkov X, ktorá sa bude rovnať h.

Nech riadok, ktorý má 0 aj 1, je i-ty v poradí a stĺpec s rovnakou vlastnosťou je j-ty v poradí. Bez ujmy na všeobecnosti, nech na ij-tom mieste matice B sa nachádza 0, teda $B_{ij} = 0$.

Potom v i-tom riadku B sa určite nachádza hodnota $B_{ik} = 1$ a v j-tom stĺpci sa nachádza hodnota $B_{lj} = 1$, pričom, samozrejme, $k \neq j$ a $l \neq i$. Označme $x_{ij}, x_{ik}, x_{lj}, x_{lk}$ riadky matice X prislúchajúce prvkom $B_{ij}, B_{ik}, B_{lj}, B_{lk}$ matice B. Potom:

$$x_{ij} - x_{ik} - x_{lj} + x_{lk} = (0, 0, 0, \ldots, 0, N_1, N_2)$$

je vektor, ktorý má na prvých $m+n$ miestach 0, a na posledných dvoch hodnoty N_1 a N_2, ktoré zatiaľ necháme bokom. Prečo má daný vektor na prvých $m+n$ miestach 0? Vo všeobecnosti má riadok x_{rs} matice X prislúchajúci prvku B_{rs} matice B na prvých $m+n$ miestach dve 1, jednu prislúchajúcu riadku, druhú stĺpcu matice B. Konkrétne, riadok x_{rs} má 1 na r-tom mieste a $(m+s)$-tom mieste.

Súčet $x_{ij} + x_{lk}$ má teda štyri jednotky na miestach $i, j, m+j, m+k$. Súčet $x_{ik} + x_{lj}$ má jednotky na tých istých miestach, preto $x_{ij} - x_{ik} - x_{lj} + x_{lk} = x_{ij} + x_{lk} - (x_{ik} + x_{lj})$

má na prvých $m+n$ miestach 0.

Takto to vyzerá, keď zanedbáme stĺpce so samými nulami:

$$
\begin{aligned}
&+(1,0,1,0)\\
&-(1,0,0,1)\\
&-(0,1,1,0)\\
&+(0,1,0,1)\\
&=(0,0,0,0)
\end{aligned}
$$

Ako vyzerá chvost (N_1, N_2)? Keďže $B_{ij} = 0$ a $B_{ik} = B_{lj} = 1$, výraz $x_{ij} - x_{ik} - x_{lj}$ nám vytvorí chvost $(1, -2)$ (prípadne $(-2, 1)$, na tom ale nezáleží). Potom na základe toho, či na mieste B_{lk} matice návrhu bola 0 alebo 1, nám výraz $x_{ij} - x_{ik} - x_{lj} + x_{lk}$ dá na posledných dvoch miestach $(2, -2)$ alebo $(1, -1)$. Vo všetkých prípadoch vektor h dostaneme ihneď, prípadne po prenásobení konštantou. Našli sme teda lineárnu kombináciu riadkov X, ktorá nám dala vektor h, preto $h \in \mathcal{M}(X^T)$. $\square$

Uvedomme si, že toto tvrdenie dáva intuitívne význam. Ak by sme totiž organizovali experiment napr. podľa druhej matice so zoznamu vyššie, tak by sme nevedeli rozlíšiť, či sú pozorované rozdiely medzi meraniami zodpovedajúcimi prvým dvom stĺpcom a posledným stĺpcom spôsobené rozdielom medzi efektami ošetrení, alebo rozdielom medzi efektami ošetrovateľov 1 a 2 v porovnaní s ošetrovateľom 3. (Bavíme sa o reprezentácii modelu, kde riadky reprezentujú pacientov a stĺpce ošetrovateľov.)

4.2 Ekvivalencie binárnych matíc

Pre rozmer $m \times n$ existuje 2^{mn} binárnych matíc, čo je obrovské číslo už pre malé hodnoty m a n. (Napríklad už pre rozmer 5×5 máme viac než milión matíc.) Pre skúmanie špecifík jednotlivých matíc je výhodné uvedomiť si niektoré očividné ekvivalencie medzi maticami, čo nám môže výrazne zredukovať počet matíc, s ktorými pracujeme, a v konečnom dôsledku nám to umožní pracovať spôsobom "brute force" do väčšieho rozmeru.

Hypotéza 4.1. *Matice, ktoré možno spermutovať jednu na druhú riadkovými a stĺpcovými permutáciami, sú ekvivalentné v zmysle, že majú rovnakú hodnotu minimálneho rozptylu Gaussovho-Markovovho odhadu rozdielu medzi efektami ošetrení.*

Hypotézu uvádzame bez matematického dôkazu, ukážeme však, že ak naša interpretácia modelu reprezentovaného binárnou maticou je správna, hypotéza musí platíť. Model sme interpretovali ako zoznam trojíc pacient, lekár, spôsob prijatia lieku, pričom pacienti zodpovedali riadkom, lekári stĺpcom, a spôsob prijatia lieku hodnotám v matici. Uvedomme si, že o lekároch ani pacientoch pri výpočte rozptylu nič nevieme. Permutáciám riadkov a stĺpcov preto zodpovedá akési „preznačenie" jednotlivých lekárov a pacientov, vôbec to ale nezmení štruktúru experimentu. Pripomeňme si, že rozptyl Gaussovho-Markovovho odhadu rozdielu medzi efektami počítame pred vykonaním experimentu a nameraním dát.

Táto jediná ekvivalencia nám výrazne zredukuje počet matíc, s ktorými musíme počítať. V tabuľke uvádzame počet neekvivalentných tried pre štvorcové rozmery až do 6×6.

Rozmer	1×1	2×2	3×3	4×4	5×5	6×6
Počet binárnych matíc	2	16	512	65 536	33 554 432	68 719 476 736
Počet neekvivalentných tried	2	7	36	317	5 624	251 610

4.3 Optimálnosť návrhu experimentu

Teraz sa pokúsime nájsť modely, ktoré pre nás budú v určitom zmysle optimálne. Najprv ale definujme, čo presne optimalita modelu znamená.

Definícia 4.2. *Nech $h = (0, 0, \ldots, +1, -1)^T$ je vektor zodpovedajúci lineárnej kombinácii dvoch efektov a nech $B \in \mathbb{R}^{m \times n}$ je taká binárna matica, že pre jej informačnú maticu M_B platí $h \in \mathcal{M}(M_B)$. Potom B nazývame optimálnou, ak pre všetky binárne matice $C \in \mathbb{R}^{m \times n}$ také, že $h \in \mathcal{M}(M_C)$ pre ich informačnú maticu M_C, platí:*

$$h^T M_B^- h \leq h^T M_C^- h$$

kde M_B^- a M_C^- sú ľubovoľné pseudoinverzie informačných matíc M_B a M_C.

Nájsť optimálne matice v zmysle definície (4.2) je cieľom tejto práce. Podobne ako pri skúmaní, či hodnotu $h^T b$ vôbec budeme môcť odhadnúť, preskúmame najprv matice malého rozmeru a na základe výsledkov vyslovíme hypotézu pre všeobecný rozmer.

Postupovať budeme nasledovne: zoberieme si malý rozmer a pre všetky binárne matice tohto rozmeru spočítame hodnotu $h^T M^- h$, pokiaľ je to možné (viď tvrdenie (4.1)). Z množiny týchto hodnôt vyberieme minimum. Optimálne návrhy budú tie, ktorých prislúchajúca hodnota bude práve toto minimum.

Optimálne návrhy sme určili do rozmeru 6×5, tu sú niektoré z nich pre štvorcové rozmery:

$$\begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 1 \end{bmatrix}$$

Ďalšie matice sa nachádzajú v prílohe B. Na základe týchto návrhov sme pre štvorcové matice prišli s nasledovnou hypotézou.

Hypotéza 4.2. *(pre štvorcové matice)*

Binárna matica návrhu B rozmeru $2k \times 2k$ je optimálna v zmysle definície (4.2) práve vtedy, keď má v každom riadku a v každom stĺpci práve k jednotiek (a práve k núl).

Binárna matica návrhu B rozmeru $(2k+1) \times (2k+1)$ je optimálna práve vtedy, keď všetky jej riadky aj stĺpce majú rovnaký počet jednotiek, a ten počet je buď k alebo $k+1$.

Poznámka: Hypotéza je do rozmeru 6×6 tvrdením, pretože sa nám to podarilo ukázať úplnou enumeráciou všetkých možných návrhov.

Ako to vyzerá pre matice iného, ako štvorcového rozmeru? Kľúčový je pomer jednotiek a núl v matici návrhu. Všimnime si, že pri štvorcových rozmeroch v predošlej hypotéze je v prípade párneho rozmeru rovnaký počet jednotiek a núl, no v prípade nepárneho rozmeru pomer nielenže nie je pol na pol (čo ani nemôže byť), ale dokonca k tomuto pomeru ani nie je tak blízko, ako by mohol byť. Napríklad pri rozmere 5×5 by sme mohli očakávať, že pomer jednotiek a núl v optimálnom návrhu bude $13 : 12$, ale podľa hypotézy je to $15 : 10$. Skutočnosť, že 5 delí 10 aj 15, nie je náhodná, a bude zohrávať kľúčovú rolu pri hypotéze o návrhoch iných ako štvorcových rozmerov.

Hypotéza 4.3. *(pre neštvorcové matice)*

Postačujúcou podmienkou optimality pre rozmer $2k \times 2l$ *je, ak má matica práve* $2kl$ *jednotiek a* $2kl$ *núl, pričom v každom z* $2k$ *riadkov je* l *jednotiek a* l *núl a v každom z* $2l$ *stĺpcov je práve* k *jednotiek a* k *núl.*

Postačujúcou podmienkou optimality pre rozmer $2k \times 2l+1$ *je, ak má matica práve* $(2l+1)k$ *jednotiek a* $(2l+1)k$ *núl, pričom v každom z* $2l+1$ *stĺpcov je práve* k *jednotiek a* k *núl, v* k *riadkoch je* $l+1$ *jednotiek a* l *núl a v* k *zvyšných riadkoch je* l *jednotiek a* $l+1$ *núl.*

Ďalej nech bez ujmy na všeobecnosti $k < l$.

Postačujúcou podmienkou optimality pre rozmer $(2k+1) \times (2l+1)$ *je, ak má matica* $(l+1)(2k+1)$ *jednotiek a* $l(2k+1)$ *núl (alebo naopak), pričom v každom z* $2k+1$ *riadkov je práve* $l+1$ *jednotiek a* l *núl, v* $l+k+1$ *stĺpcoch je práve* $k+1$ *jednotiek a v* $l-k$ *stĺpcoch je práve* k *jednotiek.*

Poznámka: Podmienky nie sú nutné, pretože napr. pri rozmere 3×6 sme našli optimálne matice s pomerom jednotiek a núl nielen $9:9$ (spĺňajúce hypotézu), ale aj $6:12$, čo naznačuje, že v niektorých prípadoch matice pokryté hypotézou nie sú všetky optimálne matice daného rozmeru. Hypotéza je napriek tomu cenná, pretože pre prax nám poskytuje mechanizmus, ako vytvoriť optimálny návrh experimentu.

Niektoré z optimálnych návrhov pre neštvorcové rozmery:

$$\begin{bmatrix} 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 \end{bmatrix}$$

Niektoré z optimálnych návrhov pre rozmer 3×6:

$$\begin{bmatrix} 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 & 0 & 0 \end{bmatrix}$$

Napriek tomu, že hypotézu o optimálnych maticiach sa nám matematicky dokázať nepodarilo, numericky sme až do rozmeru 30×30 ukázali, že zmena jednej hodnoty v nami navrhovanom optimálnom návrhu spôsobí zväčšenie hľadaného rozptylu. Dané matice teda určite tvoria lokálne minimum (na okolí tých matíc, ktoré majú odlišnú najviac jednu hodnotu).

4.4 Hodnota minimálneho rozptylu

Ako závisí hodnota Gaussovho-Markovovho odhadu rozdielu medzi efektami od rozmeru matice návrhu? Ukážeme, že tento najmenší dosiahnuteľný rozptyl má so stúpajúcim rozmerom klesajúcu tendenciu. (Čo, napokon, dáva zmysel, pretože väčší rozdiel nám ponúka viac dát na určenie rozdielu medzi dvoma efektmi.)

Hodnoty pre jednotlivé rozmery určíme tak, že pre každý rozmer zoberieme jednu maticu optimálnu v zmysle definície (4.2).

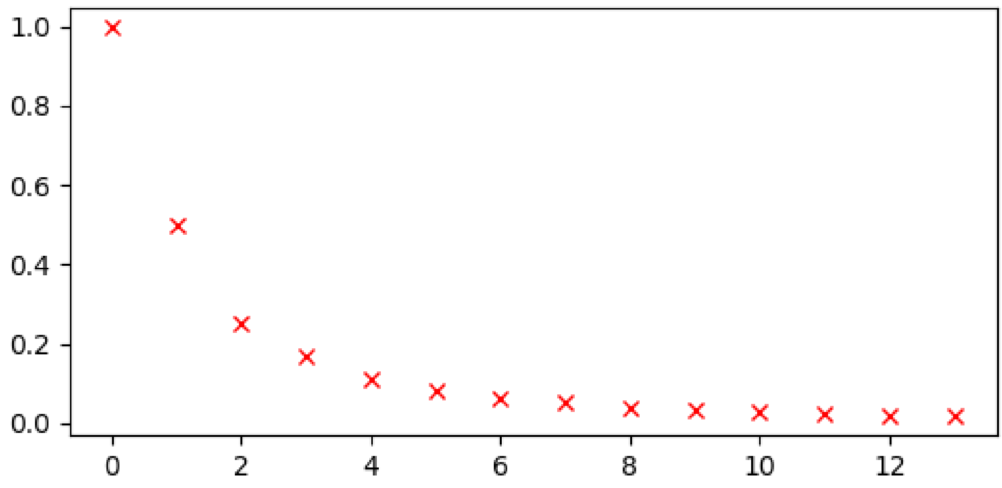
Pre párny rozmer $2k \times 2k$ rozdelíme maticu na 4 bloky $k \times k$, z nich ľavý horný a pravý dolný budú samé jednotky, zvyšné dva samé nuly. Pre nepárny rozmer $(2k+1) \times (2k+1)$ budú ľavý horný blok $k \times k$ tvoriť samé jednotky a pravý dolný blok $(k+1) \times (k+1)$ bude identita, v ktorej sa zamenili jednotky a nuly.

Názorný príklad, ako to vyzerá pre rozmery 6×6 a 7×7:

$$\begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 \end{bmatrix}$$

V tabuľke a grafe sú znázornené hodnoty rozptylu do rozmeru 15×15.

Rozmer	Minimálna hodnota rozptylu
2×2	1
3×3	0.5
4×4	0.25
5×5	0.167
6×6	0.111
7×7	0.083
8×8	0.063
9×9	0.05
10×10	0.04
11×11	0.033
12×12	0.028
13×13	0.024
14×14	0.02
15×15	0.018



Vidíme, že pre rozmer $2k \times 2k$ je hodnota rozptylu $1/k^2$. Ukážeme, že to platí vo všeobecnosti.

Tvrdenie 4.3. *Pre optimálnu maticu návrhu rozmeru $2k \times 2k$ je hodnota Gaussovho-Markovho odhadu rozdielu medzi efektami $1/k^2$.*

Dôkaz: Nech B je optimálna matica návrhu rozmeru $2k \times 2k$ spĺňajúca hypotézu (4.2). Označme X maticu lineárnej regresie pre návrh B. Hodnotu rozptylu m vypočítame ako $h^T M^- h$, kde $M = X^T X$ je informačná matica lineárnej regresie a vektor $h = (0, 0, \ldots, -1, 1)^T$ zodpovedá rozdielu dvoch efektov. Ako vyzerajú posledné dva stĺpce matice M?

Matica X je rozmeru $(4k+2) \times (4k+2)$, pričom prvých $2k$ riadkov zodpovedá $2k$ riadkom návrhu B, druhých $2k$ riadkov zodpovedá $2k$ stĺpcom návrhu B, a posledné dva riadky zodpovedajú efektu na jednotlivých pozíciách návrhu B. Posledné dva stĺpce súčinu $X^T X$ zodpovedajú násobkom riadkov X^T (stĺpcov X) s dvoma stĺpcami X zodpovedajúcimi efektu. Ak riadok, ktorým násobíme, je jeden zo $4k$ riadkov zodpovedajúcich stĺpcu alebo riadku návrhu B, hodnota súčinu bude predstavovať počet toho ktorého efektu v danom stĺpci alebo riadku. Tá je podľa hypotézy (4.2) práve k pre oba efekty. Ak navzájom násobíme riadok a stĺpec zodpovedajúci efektu, hodnota bude 0, ak ide o rozdielne efekty, alebo celkový počet daného efektu v návrhu B, ak ide o ten istý efekt. Počet efektu v optimálnom návrhu je podľa hypotézy $2k^2$. Matica $X^T X$ je symetrická, preto jej posledné dva riadky budú rovnaké ako posledné dva stĺpce.

Posledné dva riadky matice $M = X^T X$ pre maticu návrhu B rozmeru $2k \times 2k$ majú teda tvar:

$$\begin{bmatrix} k & k & \ldots & k & 2k^2 & 0 \\ k & k & \ldots & k & 0 & 2k^2 \end{bmatrix}$$

Napr. takto vyzerá matica M pre optimálne návrhy rozmerov 2 a 4:

$$\begin{bmatrix} 2 & 0 & 1 & 1 & 1 & 1 \\ 0 & 2 & 1 & 1 & 1 & 1 \\ 1 & 1 & 2 & 0 & 1 & 1 \\ 1 & 1 & 0 & 2 & 1 & 1 \\ 1 & 1 & 1 & 1 & 2 & 0 \\ 1 & 1 & 1 & 1 & 0 & 2 \end{bmatrix}, \begin{bmatrix} 4 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 2 & 2 \\ 0 & 4 & 0 & 0 & 1 & 1 & 1 & 1 & 2 & 2 \\ 0 & 0 & 4 & 0 & 1 & 1 & 1 & 1 & 2 & 2 \\ 0 & 0 & 0 & 4 & 1 & 1 & 1 & 1 & 2 & 2 \\ 1 & 1 & 1 & 1 & 4 & 0 & 0 & 0 & 2 & 2 \\ 1 & 1 & 1 & 1 & 0 & 4 & 0 & 0 & 2 & 2 \\ 1 & 1 & 1 & 1 & 0 & 0 & 4 & 0 & 2 & 2 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 & 4 & 2 & 2 \\ 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 8 & 0 \\ 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 0 & 8 \end{bmatrix}$$

Teraz dosadením ukážeme, že $h^T M^- h = (h^- M h^{-T})^-$. Ak to má platiť, musí platiť nasledovná rovnosť:

$$h^- M h^{-T} = h^- M h^{-T} (h^- M h^{-T})^- h^- M h^{-T} = h^- M h^{-T} h^T M^- h h^- M h^{-T}$$

Pre vektor v vo všeobecnosti platí, že $v^- = \frac{v^T}{v^T v}$. Pre náš vektor h to znamená, že $h^- = h^T/2 = (0, 0, \ldots, -1/2, 1/2)$. Ukážeme, že $h^{-T} h^T$ a hh^- vo vzorci vyššie dávajú tú istú maticu.

$$h^{-T} h^T = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ -1/2 \\ 1/2 \end{bmatrix} \begin{bmatrix} 0 & 0 & \ldots & -1 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 & \ldots & 0 & 0 \\ 0 & 0 & \ldots & 0 & 0 \\ \vdots & & \ddots & & \\ 0 & 0 & \ldots & 1/2 & -1/2 \\ 0 & 0 & \ldots & -1/2 & 1/2 \end{bmatrix}$$

$$hh^- = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ -1 \\ 1 \end{bmatrix} \begin{bmatrix} 0 & 0 & \ldots & -1/2 & 1/2 \end{bmatrix} = \begin{bmatrix} 0 & 0 & \ldots & 0 & 0 \\ 0 & 0 & \ldots & 0 & 0 \\ \vdots & & \ddots & & \\ 0 & 0 & \ldots & 1/2 & -1/2 \\ 0 & 0 & \ldots & -1/2 & 1/2 \end{bmatrix}$$

Označme túto maticu H. Teraz ukážeme, že $HM = MH$.

$$HM = \begin{bmatrix} 0 & 0 & \ldots & 0 & 0 \\ 0 & 0 & \ldots & 0 & 0 \\ \vdots & & \ddots & & \\ 0 & 0 & \ldots & 1/2 & -1/2 \\ 0 & 0 & \ldots & -1/2 & 1/2 \end{bmatrix} \begin{bmatrix} & & & k & k \\ & & & \vdots & \vdots \\ & & & k & k \\ k & \ldots & k & 2k^2 & 0 \\ k & \ldots & k & 0 & 2k^2 \end{bmatrix} = \begin{bmatrix} 0 & 0 & \ldots & 0 & 0 \\ 0 & 0 & \ldots & 0 & 0 \\ \vdots & & \ddots & & \\ 0 & 0 & \ldots & k^2 & -k^2 \\ 0 & 0 & \ldots & -k^2 & k^2 \end{bmatrix}$$

$$MH = \begin{bmatrix} & & & k & k \\ & & & \vdots & \vdots \\ & & & k & k \\ k & \ldots & k & 2k^2 & 0 \\ k & \ldots & k & 0 & 2k^2 \end{bmatrix} \begin{bmatrix} 0 & 0 & \ldots & 0 & 0 \\ 0 & 0 & \ldots & 0 & 0 \\ \vdots & & \ddots & & \\ 0 & 0 & \ldots & 1/2 & -1/2 \\ 0 & 0 & \ldots & -1/2 & 1/2 \end{bmatrix} = \begin{bmatrix} 0 & 0 & \ldots & 0 & 0 \\ 0 & 0 & \ldots & 0 & 0 \\ \vdots & & \ddots & & \\ 0 & 0 & \ldots & k^2 & -k^2 \\ 0 & 0 & \ldots & -k^2 & k^2 \end{bmatrix}$$

Takisto platí, že:

$$h^{-}H = \begin{bmatrix} 0 & 0 & \dots & -1/2 & 1/2 \end{bmatrix} \begin{bmatrix} 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & \dots & 0 & 0 \\ \vdots & & \ddots & & \\ 0 & 0 & \dots & 1/2 & -1/2 \\ 0 & 0 & \dots & -1/2 & 1/2 \end{bmatrix} =$$

$$\begin{bmatrix} 0 & 0 & \dots & -1/2 & 1/2 \end{bmatrix} = h^{-}$$

$$Hh^{-T} = \begin{bmatrix} 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & \dots & 0 & 0 \\ \vdots & & \ddots & & \\ 0 & 0 & \dots & 1/2 & -1/2 \\ 0 & 0 & \dots & -1/2 & 1/2 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ \vdots \\ -1/2 \\ 1/2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ -1/2 \\ 1/2 \end{bmatrix} = h^{-T}$$

Dosaďme:

$$h^{-}Mh^{-T}h^{T}M^{-}hh^{-}Mh^{-T} = h^{-}MHM^{-}HMh^{-T} = h^{-}HMM^{-}MHh^{-T} =$$
$$h^{-}HMHh^{-T} = h^{-}Mh^{-T}$$

Takže naozaj platí $h^{T}M^{-}h = (h^{-}Mh^{-T})^{-}$, resp. $(h^{T}M^{-}h)^{-} = h^{-}Mh^{-T}$. Hodnotu $h^{-}Mh^{-T}$ vieme vypočítať:

$$h^{-}Mh^{-T} = \begin{bmatrix} 0 & 0 & \dots & -1/2 & 1/2 \end{bmatrix} \begin{bmatrix} & & & k & k \\ & & & \vdots & \vdots \\ & & & k & k \\ k & \dots & k & 2k^2 & 0 \\ k & \dots & k & 0 & 2k^2 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ \vdots \\ -1/2 \\ 1/2 \end{bmatrix} =$$

$$\begin{bmatrix} 0 & 0 & \dots & -k^2 & k^2 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ \vdots \\ -1/2 \\ 1/2 \end{bmatrix} = k^2/2 + k^2/2 = k^2$$

Hodnota x^{-} pre skalár x je rovná $1/x$, preto $h^{T}M^{-}h = (h^{-}Mh^{-T})^{-} = 1/k^2$. □

Poznámka: Pri návrhoch nepárneho štvorcového rozmeru dôkaz zlyhá na rovnosti $HM = MH$. Infomačná matica M totiž pri nepárnom rozmere nemá natoľko vhodnú štruktúru ako pri párnom.

Záver

TODO

Zoznam použitej literatúry

[1] Pázman, A., Lacko, V.: *Prednášky z regresných modelov*, Vydavateľstvo UK, Bratislava, 2012, 2015

[2] http://oeis.org/A002724

Príloha A