



Univerzita Komenského v Bratislave
Fakulta matematiky, fyziky a informatiky



Mgr. Alena Bachratá

Autoreferát dizertačnej práce

Optimálne navrhovanie blokových experimentov

na získanie akademického titulu
philosophiæ doctor

v odbore doktorandského štúdia
9.1.9 Aplikovaná matematika

Bratislava 2015

Dizertačná práca bola vypracovaná v dennej forme doktorandského štúdia na Fakulte matematiky, fyziky a informatiky UK v Bratislave

Predkladateľ: Mgr. Alena Bachratá
Katedra aplikovanej matematiky a štatistiky
Fakulta matematiky, fyziky a informatiky
Univerzita Komenského v Bratislave
Mlynská dolina, 842 48 Bratislava

Školiteľ: doc. Mgr. Radoslav Harman, PhD.
Katedra aplikovanej matematiky a štatistiky
Fakulta matematiky, fyziky a informatiky
Univerzita Komenského v Bratislave

Oponenti:

Obhajoba dizertačnej práce sa koná o

pred komisiou pre obhajobu dizertačnej práce v odbore doktorandského štúdia
vymenovanou predsedom odborovej komisie

v študijnom odbore 9.1.9. aplikovaná matematika

na Fakulte matematiky, fyziky a informatiky Univerzity Komenského v Bratislave,
Mlynská dolina, 842 48 Bratislava

Predseda odborovej komisie

Prof. RNDr. Marek Fila, DrSc.

Katedra aplikovanej matematiky a štatistiky

Fakulta matematiky, fyziky a informatiky

Univerzita Komenského v Bratislave

Mlynská dolina, 842 48 Bratislava

1 Úvod

Navrhovanie experimentov slúži na optimalizáciu množstva informácie získanej z experimentu, pri daných vstupných podmienkach a ohraničeniach. Potreba optimalizácie množstva získanej informácie vznikla kvôli finančnej a časovej náročnosti niektorých experimentov. Priekopníkom v oblasti navrhovania experimentov je R. A. Fisher (viď. [13]). Túto oblasť štatistiky skúmal počas práce v jednom z najstarších poľnohospodárskych výskumných centier, takže pôvodná myšlienka navrhovania experimentov vznikla na základe riešenia problémov z praxe.

My sa zameriame na blokové návrhy a ich prepojenie s teóriou grafov. Hlavným odrazovým mostíkom sú práce R. A. Baileyovej a P. J. Camerona [4], [5]. Okrem nich sa prepojením navrhovania experimentov a teórie grafov zaoberal aj napríklad T. Tjur v práci [24]. V spomínaných prácach nájdeme viacero prístupov k reprezentovaniu návrhov experimentov pomocou grafov.

Aby sme mohli ukázať výsledky práce, potrebujeme zaviesť pojmy z teórie navrhovania experimentov (čerpáme z knihy [23]) a ukázať prepojenie blokových experimentov a teórie grafov (čerpáme z článku [5]).

1.1 Navrhovanie experimentov

Majme experiment pozostávajúci z N pokusov (meraní, pozorovaní), definovaný pomocou lineárneho regresného modelu

$$Y_i = \mathbf{f}^T(\tilde{x}_i)\theta + \varepsilon_i, \quad i = 1, \dots, N. \quad (1)$$

Jednotlivé body návrhu experimentu $\tilde{x}_1, \dots, \tilde{x}_N$ sú prvky z množiny \mathfrak{X} , ktorá reprezentuje všetky možné body návrhu experimentu. Označme počet prvkov tejto množiny $|\mathfrak{X}| = n$. V každom bode návrhu môžeme vykonať viacero pokusov, teda niektoré body návrhu sa môžu opakovať. Všetky možné body návrhu experimentu usporiadajme a očísľujme x_1, \dots, x_n . Potom množina všetkých bodov návrhu je $\mathfrak{X} = \{x_1, \dots, x_n\}$. Teda $\tilde{x}_i \in \mathfrak{X}$ je v poradí i -ty pokus experimentu, $i = 1, \dots, N$. Hodnoty Y_1, \dots, Y_N sú pozorovania v jednotlivých bodoch $\tilde{x}_1, \dots, \tilde{x}_N$. Funkcia $\mathbf{f}: \mathfrak{X} \rightarrow \mathbb{R}^m$ je funkcia priradujúca regresor $\mathbf{f}(\tilde{x}_i)$ bodu $\tilde{x}_i \in \mathfrak{X}$, $\theta \in \mathbb{R}^m$ je vektor neznámych parametrov modelu a $\varepsilon_1, \dots, \varepsilon_N$ sú nezávislé, rovnako rozdelené chyby s nulovou strednou hodnotou a konečnou nenulovou varianciou σ^2 .

Maticový zápis základného modelu (1) je $\mathbf{Y} = \mathbf{F}\theta + \varepsilon$, kde $\mathbf{Y} = (Y_1, \dots, Y_N)^T$ je vektor pozorovaní, $\mathbf{F} = (\mathbf{f}(\tilde{x}_1), \dots, \mathbf{f}(\tilde{x}_N))^T$ je matica plánu typu $N \times m$, $\theta \in \mathbb{R}^m$ je stĺpcový vektor neznámych parametrov modelu a $\varepsilon = (\varepsilon_1, \dots, \varepsilon_N)^T$ je vektor chýb so $E(\varepsilon) = \mathbf{0}_N$ a $D(\varepsilon) = \sigma^2 \mathbf{I}_N$. Pre model platí $E(\mathbf{Y}) = \mathbf{F}\theta$ a $D(\mathbf{Y}) = \sigma^2 \mathbf{I}_N$.

Informačná matica pre subsystém parametrov $\mathbf{P}^T \theta$:

V experimentoch používame dva druhy parametrov. *Významné parametre*, ktorých efekty nás pri odhadovaní zaujímajú (označíme ich τ) a *pomocné parametre*, ktoré nechceme odhadovať ale pomáhajú vytvoriť presnejší model (označíme ich β). Chceme maximalizovať informáciu o *významných parametroch*, alebo ich lineárnej kombinácii.

Zavedieme lineárny podsystém *významných parametrov* $\mathbf{P}^T \theta$, kde \mathbf{P} je *matica koeficientov* typu $m \times s$, $m \geq s$, ktorá má plnú hodnotu, a s je počet lineárnych kombinácií parametrov ktoré odhadujeme. Cieľom práce je nájsť návrh experimentu, ktorý vedie k najlepšiemu lineárnemu nevychýlenému odhadu lin. subsystému parametrov $\mathbf{P}^T \theta$.

Dá sa ukázať (vid'. [23]), že subsystém $\mathbf{P}^T \theta$ je *odhadnuteľný*, ak platí podmienka odhadnuteľnosti

$$\mathcal{M}(\mathbf{P}) \subseteq \mathcal{M}(\mathbf{M}), \quad (2)$$

kde $\mathbf{F}^T \mathbf{F} = \mathbf{M}$ je *momentová matica* experimentu, resp. informačná matica pre celý systém parametrov. Ak je splnená podmienka odhadnuteľnosti (2), tak najlepší lineárny nevychýlený odhad subsystému $\mathbf{P}^T \theta$ je

$$\hat{\theta}_p = \mathbf{P}^T \mathbf{M}^- \mathbf{F}^T \mathbf{Y},$$

ktorý navyše nezávisí od voľby pseudoinverzie matice \mathbf{M} a kovariančná matica $D(\hat{\theta}_p) = \sigma^2 \mathbf{P}^T \mathbf{M}^- \mathbf{P}$ je regulárna [23].

Ak je splnená podmienka odhadnuteľnosti 2, tak je informačná matica pre subsystém parametrov $\mathbf{P}^T \theta$ regulárna a vieme ju upraviť na tvar

$$\mathbf{C}_p(\mathbf{M}) = (\mathbf{P}^T \mathbf{M}^- \mathbf{P})^{-1}. \quad (3)$$

Informačná matica $\mathbf{C}_p(\mathbf{M})$ reprezentuje množstvo informácie, ktorú poskytuje experiment o lineárnom subsystéme parametrov $\mathbf{P}^T \theta$. Takto vieme vyjadriť aj informačnú maticu pre subsystém *významných parametrov*, s ktorou pracujeme vo zvyšku práce.

Kritériá optimality:

Na porovnávanie informačných matíc slúžia kritériá optimality. My pracuje s triedou ortogonálne invariantných kritérií.

Definícia 1.1 Nech S_+^v je množina všetkých pozitívne semidefinitných matíc rozmeru $v \times v$. Potom kritérium $\phi : S_+^v \rightarrow [0, \infty)$ spĺňajúce nasledovné vlastnosti nazývame *ortogonálne invariantné informačné kritérium*.

1. Loewnerovská izotónnosť - pre každé $\mathbf{C}_1, \mathbf{C}_2 \in S_+^v$ platí

$$\mathbf{C}_1 \preceq \mathbf{C}_2 \Rightarrow \Phi(\mathbf{C}_1) \leq \Phi(\mathbf{C}_2),$$

kde symbol \preceq reprezentuje Loewnerovské usporiadanie matíc.

2. Konkávnosť - pre každé $\mathbf{C}_1, \mathbf{C}_2 \in \mathcal{S}_+^v$ a konštantu $\alpha \in (0, 1)$ platí

$$\Phi(\alpha\mathbf{C}_1 + (1 - \alpha)\mathbf{C}_2) \geq \alpha\Phi(\mathbf{C}_1) + (1 - \alpha)\Phi(\mathbf{C}_2).$$
3. Pozitívna homogenita - pre každú maticu $\mathbf{C} \in \mathcal{S}_+^v$ a konštantu $\alpha \geq 0$ platí

$$\Phi(\alpha\mathbf{C}) = \alpha\Phi(\mathbf{C}).$$
4. Polospojitosť zhora - množiny $\{\mathbf{C} \in \mathcal{S}_+^v; \Phi(\mathbf{C}) \geq b\}$ sú uzavreté pre každé $b \in \mathbb{R}$.
5. Ortogonálna invariancia - existuje funkcia $\Phi : \mathbb{R}^v \rightarrow \mathbb{R}$ taká, že pre každú maticu $\mathbf{C} \in \mathcal{S}_+^v$ platí $\Phi(\mathbf{C}) = \tilde{\Phi}(\boldsymbol{\lambda}(\mathbf{C}))$, kde $\boldsymbol{\lambda}(\mathbf{C}) = (\lambda_1(\mathbf{C}), \dots, \lambda_v(\mathbf{C}))'$ je usporiadaný nerastúci vektor vlastných hodnôt matice \mathbf{C} .

Príkladom ortogonálne invariantných kritérií optimality sú kritériá D, A a E . Nás zaujímajú odhady subsystému parametrov $\mathbf{P}^T \theta$, takže nás zaujímajú ich parciálne verzie D_p, A_p a E_p . Ukážeme jedno z nich.

Definícia 1.2 Kritérium D_p -optimality $\Phi_{D_p}(\cdot)$. Kritérium $\Phi_{D_p} : \mathcal{S}_+^m \rightarrow \langle 0, \infty \rangle$ je zobrazenie definované

$$\Phi_{D_p}(\mathbf{M}) = (\det(\mathbf{C}_p(\mathbf{M})))^{1/\nu} = (\det(\mathbf{P}^T \mathbf{M} \mathbf{P}))^{-1/\nu},$$

pre momentové matice $\mathbf{M} \in \mathcal{S}_+^m$ spĺňajúce podm. odhadnuteľnosti $\mathcal{M}(\mathbf{P}) \subseteq \mathcal{M}(\mathbf{M})$. Pre ostatné momentové matice bude $\Phi_{D_p}(\mathbf{M}) = 0$.

Optimálny návrh experimentu:

Za exaktný návrh experimentu považujeme vektor $\xi = (\xi_1, \dots, \xi_n)'$, pričom jednotlivé komponenty $\xi_i \in \{0, 1, 2, \dots\}$ predstavujú počty pokusov vykonaných v bodoch návrhu experimentu x_1, \dots, x_n , budeme používať aj zápis $\xi_i = \xi(x_i)$.

Momentová matica nezávisí na poradí jednotlivých pokusov vykonávaných v bodoch návrhu experimentu. Preto ju pre konkrétny návrh experimentu môžeme definovať pomocou usporiadanej množiny bodov návrhu $\mathcal{X} = \{x_1, \dots, x_n\}$ a ich násobností, namiesto všetkých použitých bodov návrhu $\tilde{x}_1, \dots, \tilde{x}_N$.

Definícia 1.3 Momentová matica prislúchajúca návrhu experimentu ξ je matica typu $m \times m$

$$\mathbf{M}(\xi) = \sum_{x=1}^n \xi_i \mathbf{f}(x_i) \mathbf{f}^T(x_i). \quad (4)$$

Ak $\mathbf{M}(\xi)$ spĺňa podmienku odhadnuteľnosti (2), tak je subsystém $\mathbf{P}^T \theta$ odhadnuteľný pre návrh experimentu ξ . V práci budeme pracovať len s návrhmi, pre ktoré je subsystém $\mathbf{P}^T \theta$ odhadnuteľný.

Definícia 1.4 (Optimálny návrh) Hovoríme, že návrh experimentu ξ^* je optimálny vzhľadom na kritérium Φ , ak je návrh ξ^* prípustný a kritérium Φ nadobúda pre tento návrh maximum spomedzi všetkých prípustných návrhov $\xi \in \Xi$

$$\xi^* = \operatorname{argmax}_{\xi \in \Xi} \Phi(\mathbf{M}(\xi)). \quad (5)$$

Definícia 1.5 (Schurovsky optimálny návrh) *Návrh, ktorý je optimálny vzhľadom na množinu všetkých ortogonálne invariantných kritérií optimality, nazývame Schurovsky optimálny návrh.*

Ohraničenia na zdroje a množina prípustných návrhov:

Pri hľadaní optimálneho návrhu experimentu potrebujeme poznať množinu všetkých prípustných návrhov Ξ s ktorými pracujeme. Definujeme ju pomocou *ohraničení na zdroje* a návrhu ktorý chceme augmentovať. Bežné ohraničenie na množinu Ξ je obmedzenie na celkový počet rovnocenných pokusov $\xi_1 + \dots + \xi_n \leq N$. V praxi sa však často vyskytujú komplikovanejšie ohraničenia na zdroje (materiálne, ľudské, finančné, časové), s ktorými klasické algoritmy nevedia pracovať.

Skoro všetky takéto ohraničenia vieme zapísať v tvare *k ohraničení na zdroje* ako

$$\sum_{i=1}^n a_{ri} \xi_i \leq b_r \text{ pre všetky } r \in \{1, \dots, k\}; \quad (6)$$

kde a_{ri} je množstvo čerpania z r -tého zdroja pri vykonaní jedného pokusu v i -tom bode návrhu a b_r je limit na r -tý zdroj. Skrátený maticový zápis je $\mathbf{A}\xi \leq \mathbf{b}$. Navyše musia platiť tri prirodzené podmienky

- (P1) limity na zdroje sú konečné kladné čísla, tj. $b_1, \dots, b_k \in (0, \infty)$,
- (P2) rozšírenie návrhu nemôže znížiť množstvo čerpania pre žiaden zdroj, tj. $a_{ri} \geq 0$ pre všetky $r \in \{1, \dots, k\}$ a $i \in \{1, \dots, n\}$,
- (P3) žiaden pokus nie je úplne zadarmo a jeho opakovanie musí viesť k porušeniu niektorého ohraničenia, tj. pre všetky $i \in \{1, \dots, n\}$ existuje zdroj $r \in \{1, \dots, k\}$ taký, že $a_{ri} > 0$.

Potom skoro ľubovoľné podmienky z praxe vieme zapísať ako *ohraničenia na zdroje*.

Okrem *ohraničení na zdroje* (6) majme exaktný návrh $\xi^{(0)}$ predstavujúci prípadné vopred vykonané pokusy. Na základe návrhu $\xi^{(0)}$ a *ohraničení na zdroje* (6) definujeme množinu všetkých prípustných exaktných návrhov Ξ^{ex} ako

$$\Xi^{\text{ex}} = \{\xi \in \{0, 1, 2, \dots\}^n : \xi^{(0)} \leq \xi, \mathbf{A}\xi \leq \mathbf{b}\}.$$

V týchto množinách hľadáme optimálne návrhy experimentov. Potom môžeme optimálny návrh z prepísať na tvar

$$\xi^* = \operatorname{argmax}_{\xi \in \Xi^{\text{ex}}} \Phi(\mathbf{M}(\xi)).$$

Toto je základná teoretická úloha, ktorú riešime v celej práci.

1.2 Blokové návrhy s veľkosťou bloku dva

V práci sa zameriame na blokové experimenty s veľkosťou bloku dva. Ukážeme ako súvisia so všeobecne definovaným modelom, ako ich vieme zakresliť pomocou grafov a čo nám to prinesie.

Blokové návrhy s veľkosťou bloku dva a ich reprezentácia grafmi:

Definícia 1.6 *Blokový návrh* (v, b, B) definujeme ako dve množiny, množinu ošetrení (veľkosti v) a množinu blokov (veľkosti b), a binárnu reláciu incidencie medzi týmito množinami. Binárna relácia incidencie určuje, ktoré ošetrenia sa nachádzajú v spoločných blokoch, pričom v každom bloku sa nachádza práve B ošetrení.

Budeme používať len binárne (v rovnakom bloku sa nemôže opakovať rovnaké ošetrenie) blokové návrhy s veľkosťou bloku dva $(v, b, 2)$. Veľkosť bloku dva hovorí, že v každom bloku sa nachádzajú dve ošetrenia.

Binárne blokové experimenty s veľkosťou bloku dva vieme reprezentovať pomocou grafov. Konkrétne pomocou párovacieho grafu (z ang. *concurrency graph*). Ošetrenia predstavujú v vrcholov grafu a bloky predstavujú b hrán grafu. Ak sa dve ošetrenia nachádzajú v spoločnom bloku, tak sú vrcholy reprezentujúce príslušné ošetrenie spojené hranou reprezentujúcou príslušný blok. Vďaka používaniu binárnych návrhov vzniknutý graf neobsahuje slučky. Násobné hrany predstavujú viacnásobné použitie tých istých dvoch ošetrení v rôznych blokoch. Párovací graf návrhu ξ označíme $G(\xi)$.

Informačnú maticu pre blokový experiment s veľkosťou bloku dva, pre subsystem významných parametrov, vieme odvodiť pomocou analýzy rozptylu a vyzerá nasledovne

$$\mathbf{C}_p(\xi) = \mathbf{R}(\xi) - \frac{1}{2}\Lambda(\xi),$$

kde $\mathbf{R}(\xi)$ je matica počtu použití jednotlivých ošetrení a $\Lambda(\xi)$ je párovacia matica návrhu, ktorú dostaneme ako $\Lambda(\xi) = \mathbf{N}(\xi)\mathbf{N}(\xi)'$, kde $\mathbf{N}(\xi)$ je matica incidencie párovacieho grafu pre návrh ξ . Vzniknutá matica $\mathbf{C}_p(\xi)$ je symetrická pozitívne semidefinitná matica a súčty členov v riadkoch (respektíve v stĺpcoch) sú nulové.

Laplacián grafu:

Nech $\mathbf{R}(\xi)$ a $\Lambda(\xi)$ sú matice definované v predošlom texte. Dá sa ukázať, že matica $\mathbf{R}(\xi)$ je pre graf návrhu matica, ktorá má na diagonále stupne vrcholov a inde samé nuly. Podobne matica $\Lambda(\xi)$ je pre graf návrhu matica, ktorá má na diagonále stupne vrcholov a mimo diagonály počet hrán medzi dvojicami vrcholov. Potom Laplacián grafu G je definovaný ako $\mathbf{L}(G(\xi)) = 2\mathbf{R}(\xi) - \Lambda(\xi)$.

Zjavne platí $\mathbf{L}(\xi) = 2\mathbf{C}(\xi)$. Takže ak chceme nájsť informačnú maticu pre systém významných parametrov pre návrh experimentu ξ , tak stačí vytvoriť pre návrh ξ párovací

graf a zostrojil' k nemu Laplacián. Keďže Laplacián je dvojnásobok informačnej matice a pracujeme s pozitívne homogénnymi kritériami, tak je jedno či budeme hľadať optimálny návrh na základe Laplaciánu, alebo informačnej matice.

Kritériá optimality pre blokové návrhy a majorizácia:

Definíciu ortogonálne invariantných kritérií optimality vieme previesť na definíciu založenú na majorizácii vektorov vlastných čísel *informačných matíc* (resp. Laplaciánov).

Definícia 1.7 (Majorizácia) *Nech \mathbf{x}, \mathbf{y} sú dva vektory v \mathbb{R}^v . Hovoríme, že vektor \mathbf{x} majorizuje vektor \mathbf{y} (značíme $\mathbf{x} \prec \mathbf{y}$), ak pre ich usporiadané prvky $x_1 \geq \dots \geq x_v, y_1 \geq \dots \geq y_v$ platí*

$$a \quad \begin{aligned} i. & \quad \sum_{i=1}^k x_{v-i+1} \leq \sum_{i=1}^k y_{v-i+1}, \quad \text{pre } k = 1, \dots, v-1 \\ ii. & \quad \sum_{i=1}^v x_i = \sum_{i=1}^v y_i. \end{aligned}$$

My budeme ako vektory \mathbf{x} a \mathbf{y} používať usporiadané vektory vlastných čísel informačných matíc návrhov $\mathbf{C}(\xi)$ (resp. príslušných Laplaciánov). Druhá podmienka rovnosti *ii.* je splnená, ak majú oba návrhy rovnako veľa, rovnako veľkých blokov.

Definícia 1.8 (Schurovsky konkávna funkcia) *Hovoríme, že funkcia $f : \mathbb{R}^v \rightarrow \mathbb{R}$ je Schurovsky konkávna, ak $\mathbf{x} \prec \mathbf{y}$ implikuje $f(\mathbf{x}) \geq f(\mathbf{y})$.*

Platí, že každá symetrická, konkávna funkcia je aj *Schurovsky konkávna*. Preto sú *ortogonálne invariantné* kritériá optimality, vzhľadom na vektory vlastných čísel informačných matíc, Schurovsky konkávne - sú symetrické (nezáleží na poradí pokusov) a konkávne (vlastnosť 2. z definície). Tým pádom môžeme definíciu 1.5 previesť na nasledovný tvar.

Definícia 1.9 (Schurovsky optimálny návrh) *Návrh ξ , ktorého vektor vlastných čísel $\lambda(\xi)$ informačnej matice pre subsystem významných parametrov je majorizovaný vektorom vlastných čísel $\lambda(\zeta)$ ľubovoľného iného prípustného návrhu ζ , nazývame Schurovsky optimálny návrh (resp. najlepší návrh vzhľadom na majorizáciu).*

Nech $\mathbf{C}(\xi)$ (skrátene \mathbf{C}) je *informačná matica* návrhu experimentu pre efekty významných parametrov τ_1, \dots, τ_v a nech $\lambda_1(\mathbf{C}) \geq \dots \geq \lambda_{v-1}(\mathbf{C}) \geq \lambda_v(\mathbf{C}) = 0$ sú jej vlastné čísla. Používame tri kritériá optimality D, A a E , v teórii blokových experimentov (pozri prácu [5]). Tu ukážeme len jedno.

Definícia 1.10 Kritérium D -optimality $\Phi_D(\mathbf{C}(\xi))$ *položíme rovné v -tej odmocniny zo súčiny všetkých vlastných čísel informačnej matice, okrem najmenšieho vlastného čísla spomedzi nich, teda v -tej odmocniny zo súčiny vlastných čísel $\lambda_1(\mathbf{C}), \dots, \lambda_{v-1}(\mathbf{C})$. Ak je niektoré vlastné číslo spomedzi $\lambda_1(\mathbf{C}), \dots, \lambda_{v-1}(\mathbf{C})$ rovné nule, tak aj kritérium Φ_D je rovné nule.*

Toto kritérium vieme interpretovať v klasickej teórii navrhovania experimentov ako parciálne kritérium D_p -optimality (podobne aj pre A_p - a E_p -optimality) pre odhad akéhokoľvek ortonormálneho systému $(v - 1)$ kontrastov efektov v významných parametrov, vďaka nasledujúcej vete.

Veta 1.11 *Nech $\mathbf{Z}_{v \times (v-1)} = (\mathbf{z}_1, \dots, \mathbf{z}_{v-1})$, $\mathbf{Z}^T \mathbf{Z} = \mathbf{I}_{v-1}$, $\mathbf{Z}^T \mathbf{1}_v = \mathbf{0}_{v-1}$ je matica, ktorej stĺpce sú vektory definujúce kontrasty medzi významnými parametrami τ_1, \dots, τ_v . Nech $\mathbf{P}^T = (\mathbf{Z}^T, \mathbf{0}_{(v-1) \times b})$. Nech $\lambda_i(\cdot)$ sú vlastné hodnoty pozitívne semidefinitnej matice usporiadané od najväčšej po najmenšiu. Nech návrh ξ spĺňa podmienku odhadnuteľnosti $\mathcal{M}(\mathbf{P}) \subseteq \mathcal{M}(\mathbf{M}(\xi))$. Potom platí*

- i.
$$\mathbf{C}_p(\xi) = \mathbf{Z}^T \mathbf{C}_{\begin{pmatrix} I_v \\ 0_{b \times v} \end{pmatrix}}(\xi) \mathbf{Z}$$
- ii.
$$\lambda_i(\mathbf{C}_{\begin{pmatrix} I_v \\ 0_{b \times v} \end{pmatrix}}(\xi)) = \lambda_i(\mathbf{C}_p(\xi)), \quad i = 1, \dots, v - 1.$$

Dôkaz vety je uvedený v dizertačnej práci.

Kritériá optimality pomocou vlastností grafov:

Kritériá optimality vieme vyjadriť aj pomocou vlastností grafov. Opäť to ukážeme iba na D -optimality. Vychádzame z Kirchhoffovej vety o počte kostier grafu [21].

Veta 1.12 *Nech G je graf s v vrcholmi. Potom nasledujúce hodnoty sú rovnaké:*

- i. počet kostier grafu;
- ii. súčin vlastných čísel Laplaciánu $\lambda_1(\mathbf{L}), \dots, \lambda_{v-1}(\mathbf{L})$ vydelený počtom vrcholov v .

Parciálne kritérium D_p -optimality pre návrhy so súvislými grafmi je v -ta odmocnina zo súčinu vlastných čísel $\lambda_1(\mathbf{C}), \dots, \lambda_{v-1}(\mathbf{C})$ informačnej matice \mathbf{C} . Laplacián $\mathbf{L} = 2\mathbf{C}$, takže súčinu vlastných čísel $\lambda_1(\mathbf{C}), \dots, \lambda_{v-1}(\mathbf{C})$ je rovný v -násobku počtu kostier grafu deleno 2^{v-1} . Pri hľadaní D -optimálneho návrhu teda stačí hľadať návrh reprezentovaný grafom, ktorý má maximálny počet kostier spomedzi grafov s v vrcholmi a b hranami.

Hľadanie grafu s maximálnym počtom kostier, pri danom počte vrcholov a hran, je v teórii grafov známe ako úloha t -optimality. V tomto znení ju spomíname v druhej kapitole.

Aplikácia - mikročipové experimenty:

Ako aplikáciu uvádzame mikročipové experimenty. Viac sa o nich dá dozvedieť napríklad z prác pani R. A. Baileyovej [3], [4] a viac podrobností je v diplomovej práci [1].

2 Ciele dizertačnej práce

Hlavné ciele dizertačnej práce sme rozdelili do troch okruhov.

- **Teoretické výsledky založené na prepojení navrhovania experimentov s teóriou grafov:**

Návrh blokového experimentu, s veľkosťou bloku dva, vieme zapísať pomocou párovacieho grafu. Vďaka tomuto zápisu vieme vyjadriť informačnú maticu návrhu experimentu pomocou Laplaciánu príslušného grafu (viď. [7]). Potom vieme kritériá optimality na základe informačnej matice návrhu, previesť na kritériá optimality vychádzajúce z vlastných čísiel Laplaciánu a tie na niektoré vlastnosti grafov. Napríklad známe kritérium D -optimality, pracujúce s determinantom informačnej matice, je ekvivalentné s počtom kostier párovacieho grafu (viď. [5]). Cieľom práce je podrobnejšie sa venovať týmto prepojeniam medzi teóriou navrhovania blokových experimentov a teórie grafov. Konkrétne chceme nájsť a rozšíriť vhodné tvrdenia z teórie grafov (napr. [20], [10]), na základe ktorých získame nové triedy optimálnych návrhov experimentov. Okrem hľadania klasických optimálnych návrhov experimentov sa zameriame na dodatočnú optimálnu augmentáciu (rozširovanie) už vykonaných experimentov.

- **Algoritmy určené na hľadanie optimálnych návrhov:**

Na hľadanie optimálnych experimentov existuje veľké množstvo algoritmov a heuristík (napr. [12],[22]), takže na prvý pohľad sa zdá, že nemá zmysel hľadať ďalšie. Medzi známymi algoritmami sa ale nenachádza žiadny, ktorý by vedel univerzálne pracovať so všeobecnými *ohraničeniami na zdroje*. Buď vedia použiť len jedno ohraničenie na celkový počet pokusov (a nedajú sa ľahko rozšíriť na viac ohraničení), alebo vedieť pracovať aj s viacerými ohraničeniami, ale sú špecializované len na jedinú konkrétnu úlohu. V praxi sa naopak pomerne často vyžaduje viacero ohraničení (napr. [9, 12, 14]).

Naším druhým cieľom je nájsť vhodný algoritmus, ktorý by vedel pracovať s ľubovoľným množstvom *ohraničení na zdroje* pre čo najuniverzálnejšie zadanie. Opäť nás okrem normálneho navrhovania experimentov zaujíma aj augmentáciu už vykonaných návrhov.

- **Numerické overenie známych hypotéz:**

Posledným cieľom dizertačnej práce je overiť, pomocou nových algoritmov, niekoľko hypotéz o známych grafoch, prislúchajúcich optimálnym návrhom (napr. [10], [11], [17]). Okrem toho použiť nové algoritmy na hľadanie optimálnych návrhov experimentov pre konkrétne úlohy.

3 Výsledky dizertačnej práce

Výsledky dizertačnej práce môžeme rozdeliť do troch kategórií.

3.1 Teoretické výsledky

V teoretickej časti sme vychádzali z dvoch postupov ako pracovať s vlastnými číslami Laplaciánov párovacích grafov návrhov experimentu. Prvý z nich je založený na triviálnom rozšírení výsledkov pána Kelmansa [18, 19], konkrétne vete

Veta 3.1 (Rozšírený Kelmans) *Nech G_{v_1}, \dots, G_{v_k} sú grafy s v_1, \dots, v_k vrcholmi, pre ktoré platí $\sum_{i=1}^k v_i = v$. Potom Laplaceov polynóm pre spojenie grafov $G_{v_1} * \dots * G_{v_k}$ v bode x je*

$$\mu(G_{v_1} * \dots * G_{v_k}, x) = \frac{x(x-v)^{k-1}}{\prod_{i=1}^k (x-v+v_i)} \prod_{i=1}^k \mu(G_{v_i}, x-v+v_i).$$

Druhý postup je založený na článku pána Constantineho [10], konkrétne vetách

Veta 3.2 (Constantine 1) *Majme kompletný graf K_v na v vrchoch. Graf vytvorený odobratím k -nesusedných hrán z grafu K_v je Schurovsky optimálny.*

Veta 3.3 (Constantine 2) *Nech v je nepárne. Nech N je graf vytvorený z kompletného grafu K_v na v vrchoch, odobratím ľubovoľných $\frac{1}{2}(v-1)$ nesusedných hrán. Potom graf vytvorený z grafu N odobratím ľubovoľnej ďalšej hrany, incidentnej s vrcholom stupňa $v-1$, je Schurovsky optimálny.*

V týchto vetách sa síce pracuje s odoberaním hrán a nie s pridávaním (augmentáciou), ale vieme upraviť časť ich dôkazu tak, aby sa dali použiť aj na augmentáciu.

Na základe kombinácie týchto dvoch postupov, a ich vhodných rozšírení, sme dokázali nasledujúce tvrdenia. Našli sme optimálnu augmentáciu návrhov, ktorých párovacie grafy sú hviezdicové grafy.

Veta 3.4 (Hviezdicový graf 1) *Majme hviezdicový graf S_v na v vrchoch s $v-1$ hranami. Jednoduchý graf, vytvorený pridaním k -nesusedných hrán do grafu S_v , je Schurovsky najlepšia augmentácia grafu S_v .*

Veta 3.5 (Hviezdicový graf 2) *Nech v je nepárne. Jednoduchý graf N , vytvorený pridaním $(v-1)/2$ nesusedných hrán a ešte jednej hrany incidentnej s vrcholom stupňa $v-1$ do hviezdicového grafu S_v na v vrchoch, je Schurovsky najlepšia augmentácia grafu S_v .*

Ďalej optimálna augmentácia návrhov, ktorých párovacie grafy sú kompletné grafy.

Veta 3.6 (Kompletný graf 1) *Majme kompletný graf K_v na v vrchoch. Graf vytvorený pridaním k -nesusedných hrán do grafu K_v je Schurovsky najlepšia augmentácia grafu K_v .*

Veta 3.7 (Kompletný graf 2) *Nech v je nepárne. Nech N je graf vytvorený z kompletného grafu K_v na v vrchoch pridaním ľubovoľných $\frac{1}{2}(v-1)$ nesusedných hrán. Potom graf vytvorený z grafu N pridaním ľubovoľnej ďalšej hrany incidentnej s vrcholom stupňa $v-1$ je Schurovsky najlepšia augmentácia grafu K_v .*

Nakoniec optimálna augmentácia návrhov, ktorých párovacie grafy sú kompletne multipartitné grafy s rovnako veľkými partíciami.

Veta 3.8 (Kompletný regulárny multipartitný graf) *Majme daný kompletný multipartitný graf $K_{v_1 \times k_1}$ s k_1 rovnako veľkými partíciami, veľkosti v_1 a celkovým počtom vrcholov $v = v_1 k_1$. Jednoduchý graf H^* , vytvorený pridaním k -nesusedných hrán do ľubovoľných partícií $K_{v_1 \times k_1}$, je Schurovsky najlepšia augmentácia grafu $K_{v_1 \times k_1}$.*

3.2 Algoritmy

Druhá, dôležitá, časť výsledkov sú tri stochastické algoritmy určené na hľadanie optimálnych návrhov a optimálnych augmentácií návrhov experimentov. Prvý algoritmus je založený na princípe simulovaného žihania [6] a pracovali sme s ním už v diplomovej práci [1]. Druhý algoritmus rozdelí množinu všetkých prípustných návrhov na maximálne návrh (tj. návrhy, do ktorých nemôžeme pridať pokus do žiadneho bodu experimentu, bez porušenia niektorého z *ohraničení na zdroje*) a nemaximálne návrhy. Potom sa snaží rozumne prechádzať po maximálnych návrhoch, pretože niektorý z maximálnych návrhov je určite optimálny, vďaka monotónnosti kritérií optimality. Tento algoritmus sme predstavili v článku [2].

Tretí algoritmus je najprepracovanejší a dosahuje najlepšie výsledky nie len v porovnaní v predošlými dvomi, ale aj s inými známymi algoritmi. Podrobnosti sa dajú pozrieť v článku [15]. Hlavnou jeho výhodou je, že vie pracovať s ľubovoľným množstvom *ohraničení na zdroje*, funguje so všetkými ortogonálne invariantnými kritériami optimality (vie použiť aj ľubovoľné iné monotónne kritériá optimality). Porovnali sme ho s niekoľkými inými algoritmi, aj keď vždy len na priestore s jediným ohraničením, inak by sa iné algoritmy nedali použiť.

Okrem úloh navrhovania optimálnych experimentov vieme pre vstup do tretieho algoritmu upraviť viacero ťažkých problémov z diskkrétnej optimalizácie a kombinatoriky. Napríklad takzvaný knapsack problém (viď. [16]), teóriu spoľahlivosť počítačových sietí (viď. [8]), a hľadanie t -optimálnych grafov v teórii grafov (viď. [20]).

Na vysvetlenie princípu algoritmu, potrebujeme definovať niekoľko pojmov. Nech Ξ^{ex} je množina všetkých prípustných exaktných návrhov. Hovoríme, že n -rozmerný návrh $\zeta \in \Xi^{\text{ex}}$ je vytvorený z návrhu $\xi \in \Xi^{\text{ex}}$ *krokom dopredu* (resp. *krokom dozadu*) ak platí $\zeta = \xi + e_i$ (resp. $\zeta = \xi - e_i$), kde $e_i \in \mathbb{R}^n$ je n -rozmerný vektor samých núl s jednotkou na i -tej pozícii. Ďalej zavedieme pojem *horných susedov* návrhu $\xi \in \Xi^{\text{ex}}$, ako množinu všetkých prípustných návrhov, do ktorých sa vieme dostať z ξ pomocou jedného *kroku dopredu*. Takže množina *horných susedov* návrhu ξ je $\mathcal{U}(\xi) = \{\xi + e_1, \dots, \xi + e_n\} \cap \Xi^{\text{ex}}$. Analogicky je množina *dolných susedov* návrhu ξ rovná $\mathcal{L}(\xi) = \{\xi - e_1, \dots, \xi - e_n\} \cap \Xi^{\text{ex}}$.

Vďaka podmienkam na *ohraničenia na zdroje* platí, že ľubovoľný prípustný exaktný návrh je dostupný z ľubovoľného iného prípustného exaktného návrhu len pomocou *krokov dopredu* a *krokov dozadu* po množine Ξ^{ex} . Rovnako ako pri druhom algoritme aj tu platí, že optimálny návrh je jeden z maximálnych návrhov vďaka monotónnosti kritéria.

Priebeh algoritmu:

Algoritmus začína v počiatočnom návrhu a vykonáva exkurzie po množine prípustných exaktných návrhov. Počas exkurzií využíva *tabu list* V *charakteristických atribútov* návrhov $\text{attr}(\xi)$, ktoré už navštívil a snaží sa do nich nevracať. Zavedieme pojem *lokálneho odhadu významu* návrhu $\text{val}(\xi)$, ako reálne číslo odhadujúce ako veľmi sľubný je návrh ξ ako súčasť exkurzie vedúcej k optimálnemu návrhu. Tieto parametre sú presne definované v práci.

Nech ξ prezentuje aktuálny návrh v ktorom sa počas exkurzie nachádzame. Algoritmus najskôr skúsi vykonať *krok dopredu* (ak sa *atribút* aktuálneho návrhu nenachádza na *tabu liste* $\text{attr}(\xi) \notin V$) alebo *krok dozadu* (ak sa *atribút* aktuálneho návrhu nachádza na *tabu liste* $\text{attr}(\xi) \in V$). Týmto krokom sa presunieme do susedného prípustného exaktného návrhu ζ . Pričom návrh ζ vyberáme tak, že má maximálnu hodnotu $\text{val}(\zeta)$ spomedzi všetkých susedných návrhov ktoré nie sú na *tabu liste* $\text{attr}(\zeta) \notin V$.

Ďalej ak sa algoritmus pokúsi o *krok dopredu* ale neexistuje žiaden *horný sused* $\zeta \in \mathcal{U}(\xi)$ ktorého *atribút* by nebol na *tabu liste* $\text{attr}(\zeta) \notin V$, alebo naopak ak sa algoritmus pokúsi o *krok dozadu* ale neexistuje žiaden *dolný sused* $\zeta \in \mathcal{L}(\xi)$ ktorého *atribút* by nebol na *tabu liste* $\text{attr}(\zeta) \notin V$, tak sa algoritmus pokúsi urobiť *krok* opačným smerom ako chcel pôvodne. Ak ani potom nenájde žiaden návrh do ktorého môže prejsť, tak túto patovú situáciu vyrieši náhodným prechodom do ľubovoľného prípustného susedného návrhu $\zeta \in \mathcal{L}(\xi) \cup \mathcal{U}(\xi)$ bez ohľadu na *tabu list*.

Vždy keď algoritmus navštívi *maximálny* návrh, tak overí či tento návrh nie je lepší ako doteraz najlepší dosiahnutý návrh. Ak je nový návrh lepší, tak pokračujeme ďalej s ním. Ak počet *krokov dozadu* prekročí danú hranicu, tak exkurziu prehlásime za neúspešnú a začneme novú exkurziu z doteraz najlepšieho nájdeneho návrhu. Aj neúspešné

exkurzie sa zapisujú do *tabu listu*, tým pádom ďalšia exkurzia je iná ako predošlá. Navyše smeruje k ešte nenavštíveným návrhom. Algoritmus skončí po vopred zadanom čase a vráti najlepší dovtedy nájdený návrh.

3.3 Numerické výsledky

V poslednej časti predstavíme numerické výsledky získané pomocou tretieho algoritmu. Algoritmus sme spúšťali na rôzne dlhé časy, podľa riešených úloh. Keď sme hľadalo optimálny návrh pre úlohu s veľkým počtom ošetrení, alebo keď sme na základe výsledkov chceli nájsť protipríklad k nejakej hypotéze, tak sme ho spúšťali na hodinu a dlhšie. Pri jednoduchších úlohách, kde stačilo nájsť dostatočne eficientný návrh, alebo sme dokázali nájsť optimálny návrh dostatočne rýchlo, sme ho púšťali na kratšie.

Silno regulárne grafy:

Prvá otázka, ktorú sme zodpovedali pomocou výsledkov získaných z tretieho algoritmu, je hypotéza dva z článku [11].

Hypotéza 3.9 (Silno regulárne grafy) *Majú súvislé, silno regulárne grafy maximálnu komplexitu?*

Komplexita grafu znamená počet kostier grafu. Graf s maximálnou komplexitou je t -optimálny, resp. D -optimálny. Zaujímali sme sa o silno regulárne grafy (srg) na 16 vrcholoch. Pre tri z nich, konkrétne Shrikhande graf - $srg(16, 6, 2, 2)$ s 48 hranami; komplement k Shrikhande grafu - $srg(16, 9, 4, 6)$ s 72 hranami; komplement k Clebschovmu grafu - $srg(16, 5, 0, 2)$ s 80 hranami, sme našli na rovnakom počte hrán a vrcholov grafy s väčším počtom kostier. Pričom hodnoty kritérií D -optimality návrhov prislúchajúcich k týmto trom silno regulárnym grafom sú postupne 98.65%, 99.68% a 99.61% z hodnôt kritérií D -optimality optimálnych návrhov nájdených tretím algoritmom. Takže túto hypotézu sa pomocou výsledkov tretieho algoritmu podarilo vyvrátiť.

Ďalšie špeciálne triedy grafov: Pre žiadnu ďalšiu špeciálnu triedu grafov, sme už ďalšie hypotézy nevyvrátili. Naopak sme niektoré hypotézy empiricky overili pre menšie grafy, alebo špeciálne prípady (kompletnou enumeráciou a tretím algoritmom). Konkrétne sa jedná o nasledujúce hypotézy (nie sú empiricky potvrdené v celom znení, ale len na menších prípadoch).

Hypotéza 3.10 (Mooreove grafy) *Mooreove grafy majú maximálny počet kostier medzi všetkých jednoduchých grafov s rovnakým počtom vrcholov a hrán (viď. [17]).*

Hypotéza 3.11 (Jednoduché grafy) *Nech $G(v, e)$ je graf na v vrcholoch s e hranami, pričom $e \leq v(v - 1)/2$. Potom graf prislúchajúci k Schurovsky optimálnemu návrhu je jednoduchý.*

Ako ďalšie podporíme tvrdenie, ktoré vyslovil Constantine na konci článku [10].

Tvrdenie 3.12 (Constantine 1983) *Pri odobratí viac ako $(v+1)/2$ hrán z kompletného grafu G na v vrchoch nemôžeme byť návrh, ktorému prislúcha výsledný párovací graf Schurovsky optimálny (optimálny vzhľadom na všetky ortogonálne kritériá optimality).*

O ďalšej hypotéze si myslíme, že pôjde dokázať podobnými technikami ako predošlé uvedené teoretické výsledky v našej práci. Zatiaľ sme ju však overili iba numericky.

Hypotéza 3.13 (Kompletný skoro regulárny multipartitný graf) *Jednoduchý graf, ktorý vznikol zo skoro-regulárneho kompletného multipartitného grafu pridaním k –nesusedných hrán (pre v nepárne ešte aj s pridaním $(v-1)/2$ nesusedných hrán a jednej hrany, ktorú pridáme k vrcholu stupňa nula) je Schurovsky najlepšia augmentácia skoro-regulárneho kompletného multipartitného grafu.*

Okrem toho sme ešte tretí algoritmus porovnali so simulovaným žíhaním [25]. Tretí algoritmus našiel (až na malé výnimky) rovnaké, alebo lepšie návrhy ako simulované žíhanie. Podobne sme ukázali využitie druhého algoritmu na konkrétnej úlohe s kontrolným ošetrením. Kde kontrolného ošetrenie môžeme využiť ľubovoľné množstvo, zatiaľ čo ostatných ošetrení máme len určitý počet.

3.4 Záver

V práci sme predstavili dva druhy výsledkov. Najskôr teoretické, založené na vlastnostiach vlastných čísel Laplaciánov párovacích grafov, ktoré prislúchajú návrhom experimentu. Ako druhé prakticky použiteľný algoritmus, ktorý je použiteľný na širokú triedu problémov. Nie len na blokové experimenty s veľkosťou bloku dva. Na záver sme ukázali výsledky, získané pomocou tohto algoritmu. Myslíme si, že oblasť prepájania teórie navrhovania experimentov a teórie grafov je veľmi zaujímavá a že sa v nej nachádza ešte množstvo nevyriešených problémov.

Literatúra

- [1] Bachratá, A.: Návrh mikročipových experimentov, diplomová práca (2011)
- [2] Bachratá, A., Harman, R.: A stochastic optimization method for constructing optimal block designs with linear constraints. Proceedings from the European Young Statisticians Meeting, Osijek, 2013 (2014)
- [3] Bailey, R.A.: Design of Comparative Experiments. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press (2008)
- [4] Bailey, R.A.: Designs for two-colour microarray experiments. *J. Roy. Stat. Soc. C Appl. Stat.* 56(4), 365-394 (2007)
- [5] Bailey, R.A., Cameron, P.J.: Combinatorics of optimal designs. *Surveys in Combinatorics - London Mathematical Society Lecture Note Series 365*, 19-73 (2009)
- [6] Černý, V.: A thermodynamical approach to the travelling salesman problem: an efficient simulation algorithm. *Jurnal of Optim, Theory and Applications* 45, pages 41-51, (1985)
- [7] Cheng, C.S.: Maximizing the total number of spanning trees in a graph: two related problems in graph theory and optimum design theory. *J. Combin. Theor. B* 31, 240-248 (1981)
- [8] Chern, M.S.: On the computational complexity of reliability redundancy allocation in a series system. *Oper. Res. Lett.* 11, 309-315 (1992)
- [9] Cook, R.D., Thibodeau, L.A.: Marginally restricted D-optimal designs. *J. Am. Stat. Assoc.* 75(370), 366-371 (1980)
- [10] Constantine, G. M.: Schur convex functions on the spectra of graphs, *Discrete Mathematics* 45: 181-188, (1983)
- [11] Constantine, G. M.: Graph complexity and the Laplacian matrix in blocked experiments, *Linear and Multilinear Algebra* 28: 49-56, (1990)
- [12] Fedorov, V.V.: Optimal design with bounded density: Optimization algorithms of the exchange type. *J. Stat. Plann. Infer.* 22, 1-13 (1982)
- [13] Fisher, R. A.: *The design of Experiments*, Edinburg: Oliver and Boyd (1935)
- [14] Harman, R.: Multiplicative Methods for Computing D-Optimal Stratified Designs of Experiments. *J. Stat. Plann. Infer.* 146, 82 - 94 (2014)
- [15] Harman, R., Bachratá, A. and Filová, L.: Construction of efficient experimental designs under multiple resource constraints. *Appl. Stochastic Models Bus. Ind.*, doi: 10.1002/asmb.2117. (2015)
- [16] Hochbaum, D.S.: A nonlinear Knapsack problem. *Oper. Res. Lett.* 17(3), 103 - 110 (1995)
- [17] Jakobson, D., Rivin, I.: On some extremal problems in graph theory. *arXiv:math.CO/9907050*.(1999)

- [18] Kelmans, A. K.: The number of trees of a graph I. (Russian) *Avtomat. Telemekh.* 26: 2194-2204, (1965)
- [19] Kelmans, A. K.: The number of trees of a graph II. (Russian) *Automat. Remote Control* 27: 56-65, (1966)
- [20] Kelmans, A. K., Chelnokov, V. M.: A certain polynomial of a Graph and Graphs with an Extremal Number of Trees. *J. Combin. theory (B)* 16: 197-214, (1974)
- [21] Kirchhoff, G.: 'Über die Auflösung der Gleichungen, auf welche man bei der Untersuchung der linearen Verteilung galvanischer Ströme geführt wird. *Annals of Physical Chemistry*, 72, 497–508 (1847)
- [22] Mitchell, T.J.: An Algorithm for the Construction of "D-Optimal" Experimental Designs. *Technometrics* 16(2), 203-210 (1974)
- [23] Pukelsheim, F.: *Optimal Design of Experiments*. SIAM (2006)
- [24] Tjur, T.: Block designs and electrical networks, *Annals of Statistic* 19, 1010-1027 (1991)
- [25] Wit, E., Nobile, A., Khanin, R.: Near-optimal designs for dual-channel microarray studies. *Jurnal of the Royal Stat. Society, Series C (Applied Statistics)*, Volume 54, Issue 5, pages 817-830, (2005)

Vedecká aktivita

Zoznam publikácií

- HARMAN, R., BACHRATÁ, A. (30%), FILOVÁ, L. (2015): *Construction of efficient experimental designs under multiple resource constraints*, Applied Stochastic Models in Business and Industry [elektronický zdroj]. <http://onlinelibrary.wiley.com/doi/10.1002/asmb.2117/epdf>
ADM – Vedecké práce v zahraničných časopisoch registrovaných v databázach Web of Science alebo SCOPUS
- BACHRATÁ, A. (50%), HARMAN, R. (2013): *A stochastic optimization method for constructing optimal block designs with linear constraints*, 18th European Young Statisticians Meeting. - Osijek : J. J. Strossmayer University, 2014. - S. 147-151. - ISBN 978-953-6931-70-5
BEE – Odborné práce v zahraničných zborníkoch (konferenčných aj nekonferenčných)

Prezentácie na konferenciách

- INTERNATIONAL CONFERENCE ON TRENDS AND PERSPECTIVES IN LINEAR STATISTICAL INFERENCE: Linköping, Sweden, (2014). *Experimental Designs under Resource Constraints: Algorithmic Construction*, pozvaná prednáška.
- 18TH EUROPEAN YOUNG STATISTICIANS MEETING: J.J. Strossmayer University of Osijek, Croatia, (2013). *A stochastic optimization method for constructing optimal block designs with linear constraints*, prednáška.
- THE INTERNATIONAL CONFERENCE ON TRENDS AND PERSPECTIVES IN LINEAR STATISTICAL INFERENCE: Będlewo, Poland, (2012). *Using methods of stochastic optimization for constructing optimal experimental designs with cost constraints*, poster.

Zoznam riešených grantov

- *Metódy optimálneho navrhovania experimentov (spoluriešiteľ)*, VEGA 1/0163/13
Hlavný riešiteľ: doc. Mgr. Radoslav Harman, PhD.
- *Nové metódy matematickej štatistiky (spoluriešiteľ)*, VEGA 2/0038/12
Hlavný riešiteľ: doc. RNDr. František Rublík, CSc.
- *Navrhovanie optimálnych experimentov*, Grant UK č. UK/556/2014
- *Navrhovanie optimálnych experimentov*, Grant UK č. UK/654/2013
- *Navrhovanie blokových experimentov*, Grant UK č. UK/207/2012

Optimal design of block experiments

The main topic of the dissertation thesis is the optimal design of block experiments. We are looking for such design, that maximizes amount of information for unknown parameters of interest of linear regression model. The optimal design must be from the set of all feasible designs, which are determined by constraints on resources. We work with a big class of optimality criteria - with orthogonally invariant information criteria. We focus on block experiments with the block size two, which can be represented by concurrence graphs.

There are two kinds of results. Firstly, we show some theoretically derived classes of Schur-optimal augmentation of designs, found by applying results from graph theory. These classes are the optimal augmentations of designs, which have following concurrence graphs: Star graphs, Complete graphs, and Complete regular multipartite graphs.

Secondly, we present three novel stochastic algorithms for finding optimal exact designs with respect to the set of resource constraints. In the end of the thesis, we show numerical results of these algorithms and propose hypotheses based on the results.