ФАКУЛТЕТ ПО МАТЕМАТИКА И ИНФОРМАТИКА

Дизайни в Хемингови пространства и методи за изследване на тяхната структура

доц. д-р Христина Кулина

Семинар на ФМИ, Пампорово, ноември 23 – 25, 2016 г.

- конформационно и функционално охарактеризиране на взаимодействията на аглутинина WGA
- метал-порфиринови комплекси
- неметал-съдържащи порфирини
- WGA има регистрирана порфирин-свърваща активност
- един пофирин-свързващ център
- Изследваните метал-порфиринови съединения могат да намерят потенциално приложение в таргетната терапия на рака.
- сравняване свързващата активност на WGA
- метал-порфиринови комплекси (Pt, два Au порфирина)

- химиотерапевтичното съединение на цисплатина
- афинитетът на WGA към Au и Pt порфирини е с около два порядъка по-висок от този към цисплатина
- специализиран софтуер GraphPad Prism 5
- Институт по молекулярна биология "Акад. Румен Цанев"на БАН
- **четири** публикации, **общ IF: 4,868**
  - Molecular BioSystems
  - Biotechnology and Biotechnological Equipment

- Приложение на метода за класификационни и регресионни дървета
  - в областта на животновъдството за изследване на възможността за прогнозиране на млечността на крави в зависимост от голям брой показатели
- Използване на метода на главните компоненти и регресия с главни компоненти
  - за установяване зависимостта на усреднената млечност на крави от набор разглеждани показатели
  - при изследване на общото замърсяване на въздуха на град Димитровград

- Приложение на GPS метода за построяване на регресионен модел на CO (въглероден оксид) за град Димитровград.
- Приложение на вариационен анализ, алтернативен анализ, непараметрични и класификационни методи
  - за изследване нагласата на юноши към спортна дейност и влиянието на семейната среда за формиране на здравното им поведение
  - за проучване на разпространението, развитието, рисковите фактори и информираността на лекарите по дентална медицина за диагностицирането и лечението на моларно-инцизивна хипоминерализация (МИХ) в България

- **осем публикации**
    - ФМИ при ПУ "Паисий Хилендарски"
    - Медицински университет - гр. Пловдив
    - Тракийски университет - гр. Стара Загора
- AIP Conference Proceedings, **JSR: 0,152**
- Agricultural science and Technology
- Scientific research of the Union of Scientists in Bulgaria-Plovdiv
- Осма национална конференция "Образованието и изследванията в информационното общество"

- **осем публикации**
  - Институт по математика и информатика на БАН
  - ФМИ при СУ "Св. Климент Охридски"

- **общ IF: 1,793**
  - Designs, codes and cryptography
  - Problems of Information Transmission
  - International Workshop Algebraic and combinatorial coding theory
  - Proceedings International Workshop Optimal codes and related topics

- $Q = \{0, 1, \ldots, s-1\}$ е азбука от s символа.
- $H(n, s) = \{(x_1, x_2, \ldots, x_n) : x_i \in Q, i = 1, 2, \ldots, n\}$
- $d(x, y) = |\{i | x_i \neq y_i\}|$, $x, y \in H(n, s)$
- Всяко крайно непразно подмножество $C \subset H(n, s)$ се нарича **код**.
- Код $C \subset H(n, s)$, за който всяка $M \times \tau$ подматрица на $M \times n$ матрицата на C съдържа всяка наредена τ-орка от $H(\tau, s)$ точно $\frac{|C|}{s^\tau}$ пъти като редове, се нарича **τ-дизайн** в $H(n, s)$.
- $M = |C|$ – **мощност**; n – **размерност**.
- числото $\lambda = \frac{|C|}{s^\tau}$ се нарича **индекс** на C
- най-голямото число τ, за което C е τ-дизайн се нарича **сила на дизайна**.

- Означаваме OA (M, n, s, τ).
- $C \subset H(n, s)$ линеен код, $\tau(C) = d'(C) - 1$
- τ-дизайните в $H(n, s)$ се наричат още ортогонални масиви със сила τ или τ- независими множества.

Калямпуди Радхакришна Рао

- роден 1920 г., Индия
- 1941 –1980 г., Индийски статистически институт
- 1946 – 1948 г. Кембриджски университет

- 1948 г. – защитава докторска дисертация под ръководството на Роналд Айлмер Фишер
- 1965 г. – Кембриджския университет му присъжда престижната степен "доктор на науките"
- 1945 г. – излиза първата му книга
    - формулира теоремата на Рао–Блекуел–Колмогоров
    - формулира знаменитото неравенството на Рао-Крамер
    - въвежда Фишер-Рао метрика, Рао-разстояние
- 1949 – 1951 в поредица от доклади и статии Рао прави забележителен принос в планирането на експеримента, използвайки комбинаторни структури – ортогонални масиви.

- *On a Class of Arrangements*, Proceedings of the Edinburgh Mathematical Society, Volume 8, Issue 3, December 1949, pp. 119-125.
- Forbes Magazine – "нова мантра"за индустрията
- Приноси:
  - Оценка на статистическите параметри
  - Дисперсионен анализ
  - Комбинаторни дизайни
  - Функционални уравнения
- автор на 14 книги и 475 научни статии
- удостоен с 37 титли "доктор хонорис кауза"на 19 страни, 6 континента

- Двоичен ОА (12,11,2,2).

$$
\begin{array}{ccccccccccc}
\color{red}{0} & 0 & \color{red}{0} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\color{red}{1} & 1 & \color{red}{1} & 0 & 1 & 1 & 0 & 1 & 0 & 0 & 0 \\
\color{red}{0} & 1 & \color{red}{1} & 1 & 0 & 1 & 1 & 0 & 1 & 0 & 0 \\
\color{red}{0} & 0 & \color{red}{1} & 1 & 1 & 0 & 1 & 1 & 0 & 1 & 0 \\
\color{red}{0} & 0 & \color{red}{0} & 1 & 1 & 1 & 0 & 1 & 1 & 0 & 1 \\
\color{red}{1} & 0 & \color{red}{0} & 0 & 1 & 1 & 1 & 0 & 1 & 1 & 0 \\
\color{red}{0} & 1 & \color{red}{0} & 0 & 0 & 1 & 1 & 1 & 0 & 1 & 1 \\
\color{red}{1} & 0 & \color{red}{1} & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 1 \\
\color{red}{1} & 1 & \color{red}{0} & 1 & 0 & 0 & 0 & 1 & 1 & 1 & 0 \\
\color{red}{0} & 1 & \color{red}{1} & 0 & 1 & 0 & 0 & 0 & 1 & 1 & 1 \\
\color{red}{1} & 0 & \color{red}{1} & 1 & 0 & 1 & 0 & 0 & 0 & 1 & 1 \\
\color{red}{1} & 1 & \color{red}{0} & 1 & 1 & 0 & 1 & 0 & 0 & 0 & 1
\end{array}
$$

- $\{0,0\}$, $\{0,1\}$,$\{1,0\}$ и $\{1,1\}$

- ортогоналните масиви се използват в производствените експерименти за изучаване на ефекта от няколко фактора
- два фактора се наричат ортогонални, ако всички възможни комбинации на техните нива се срещат еднакъв брой пъти
- гарантират тестването на всички комбинации на избраните променливи
- създават множество с много по-малко тестови случаи
- тестовите случаи се определят лесно

- брой на факторите – брой стълбове на $OA(M, \mathbf{n}, s, \tau)$
- брой нива във всеки фактор – азбука на $OA(M, n, \mathbf{s}, \tau)$
- смесени (асиметрични) ортогонални масиви
- сила на дизайна – $OA(n, k, s, \tau)$
- минималното M, за което съществува $OA(\mathbf{M}, n, s, \tau)$

Основна задача

За фиксирани n, s и τ
да се намери или оцени величината $L(n, s, \tau) = \min\{\lambda : \text{съществува } (\lambda s^{\tau}, n, s, \tau) \text{ дизайн } C \subset H(n, s)\}$.

- пълен факторен факторен експеримент – $3^4 = 81$ на брой ексксперименти
- чрез OA (метод на Тагучи) използваме OA (9,4,3,2) - ще се проведат 9 експеримента.

A	**B**	**C**	**D**
A_1	B_1	C_1	D_1
A_2	B_2	C_2	D_2
A_3	B_3	C_3	D_3

0	0	0	0
0	1	1	1
0	2	2	2
1	0	1	2
1	1	2	0
1	2	0	1
2	0	2	1
2	1	0	2
2	2	1	0

Геничи Тагучи
(1924 – 2012)

- японски статистик
- награда Деминг
- почетен член на Американската асоциация за качество

- Тагучи развива идеята за използване на ортогонални масиви при планиране на експеримента и контрола върху качеството.
- Хамурапи 1792-1750 г. пр. Хр.

- развитие на методите за научно управление и осигуряване на качеството
- Фридерик Тейлър – известен като *бащата на управленската наука*, в основата на която е идеята за създаването на работен стандарт за систематизиране на производствената дейност в организацията
- Уолтър Шухарт – пръв започва да използва статистическите методи при контрола на качеството и създава известната и днес "контролна карта"
- Уилям Деминг: *Същността е в качеството! Но напразно търсите качеството между машините – то се ражда в коридорите на властта.*

- Деминг – първият лектор по статистически методи за контрол на качеството, Япония
- 50-те години на XX в., Япония
- Тагучи – три големи достижения, системата за тотално управление на качеството (СТУК)
- иновациите в областта на статистическите техники за осъществяване на експерименти
- инженеринг на качеството
- стандарт ISO 9000:2000
- извънпроизводствен контрол на качеството

- Tatsumi, K., *Test - Case - Design Support System*, Proc. of the International Conference on Quality Control (ICQC), Tokyo, 1987.
- pairwise testing (combinatorial testing)
- Pairwise Testing Available Tools: http://www.pairwise.org/tools.asp
- тестване на потребителския интерфейс
- системни тестове
- регресионни тестове
- конфигурационни тестове
- тестване на производителност

- $s = 2$, τ-дизайн $C \in H(n,2)$ с мощност $M = |C|$ и размерност n означаваме $OA(M, n; \tau)$.
- $L(n,\tau) = \min\{\lambda : \exists\ (\lambda 2^\tau, n; \tau)$ дизайн $C \subset H(n,2)\}$.
- граници на Rao, Delsarte

$$|C| \geq D(n,\tau) = \begin{cases} \sum_{j=0}^{k} \binom{n}{j}, & \text{при } \tau = 2k, \\ 2\sum_{j=0}^{k} \binom{n-1}{j}, & \text{при } \tau = 2k-1. \end{cases} \qquad (1)$$

Дизайни, достигащи границата (1) се наричат плътни.

- Bose, Bush (1952) за $\tau = 2, 3$;
- Hedayat, Stufken (1989) за $\lambda = 1$;
- Seiden, Zemash (1966)
- за $\tau = 2, 3$ – точни стойности на $L(n,\tau)$ за всяко n.

Дизайни в $H(n, 2)$

Долни граници за $L(n, \tau)$

A. S. Hedayat, N. J. A. Sloane, John Stufken, *Orthogonal arrays: Theory and Applications*, 1999.

- за всяко $4 \leq n \leq 14$ и всяко $4 \leq \tau \leq 10$.

Таблица 1.[12.1] Долни граници за $L(n, \tau)$

n \ τ	4	5	6	7	8	9	10
4	1						
5	1	1					
6	2	1	1				
7	$^{sz}4$	2	1	1			
8	4^{c}	$^{sz}4$	2	1	1		
9	6,7,8	4^{c}	4	2	1	1	
10	6,7,8	6,7,8	$^{sz}6$,7,8	4	2	1	1
11	6,7,8	6,7,8	8^{c}	$^{sz}6$,7,8	4	2	1
12	7,8	6,7,8	12–15,16	8^{c}	6,7,8	4	2
13	8	7,8	16	12–15,16	10–15,16	6,7,8	4
14	8	8	16	16	16^{c}	10–15,16	6,7,8

Levenshtein, V.I., *Universal Bounds for Codes and Designs*, Handbook of Coding Theory, Pless, V.S. and Huffman, W.C., Eds., Amsterdam: Elsevier, 1998, vol. I, ch. 6, pp. 499-648.

- $H(n,2)$ – АПМП, $\langle x,y\rangle = 1 - \frac{2d(x,y)}{n}$
- $t_i = 1 - \frac{2i}{n}, i = 0,1,\dots,n.$
- $C : OA(M,n;\tau) \subset H(n,2)$
- $y \in H(n,2),\ i = 0,1,\dots,n,$
  $p_i(y) = |\{x \in C : d(y,x) = i\}|$
- $p(y) = (p_0(y),\dots,p_n(y))$ – **спектър** на C относно y.
- $\mathbf{0} = (0,0,\dots,0) \in C$
- спектъра на $\mathbf{0}$ в C се нарича теглови спектър на C
- $w = (w_0,w_1,w_2,\dots,w_n),\ w_i = p_i(\mathbf{0}).$

- Код $C \subset H(n, 2)$ е τ-дизайн тогава и само тогава, когато за всеки полином $f(t)$ с реални коефициенти от степен най-много τ и за всяка точка $y \in H(n, 2)$ е в сила равенството

$$\sum_{x \in C} f(\langle x, y \rangle) = f_0 |C|, \tag{2}$$

където f_0 е първият коефициент в разлагането $f(t) = \sum_{i=0}^{n} f_i Q_i^{(n)}(t)$, $Q_i^{(n)}(t)$ са нормализираните полиноми на Кравчук:

$$Q_i^{(n)}(t) = \frac{1}{\binom{n}{i}} \sum_{j=0}^{i} (-1)^j \binom{d}{j} \binom{n-d}{i-j}, \quad i = 0, 1, \ldots, n.$$

Теорема. Нека $C \subset H(n,2)$ е $(M,n;\tau)$ ОА и $y \in H(n,2)$ е фиксирано. Тогава за спектъра на C относно y са в сила равенствата

$$\sum_{i=0}^{n} p_i(y)\left(1-\frac{2i}{n}\right)^k = b_k|C|, \quad k=0,1,\ldots,\tau, \qquad (3)$$

където b_k е първият коефициент (т.е. коефициента f_0) в разлагането по Кравчук на полиномите t^k.

- (3) е системата от Вандермондов тип
- $\tau+1$ уравнения с $n+1$ цели, неотрицателни неизвестни $p_0(y), p_1(y), \ldots, p_n(y)$
- $n-\tau$ свободни променливи
- $p_0(y) \geq 1,\ y \in C;\ p_0(y)=0,\ y \notin C$

За произволно избран стълб на C: $OA(M, n; \tau)$ пренареждаме редовете на дизайна, така че нулите в избрания стълб да са преди единиците.

- $C^{'} : (M, n-1; \tau)$
- $C_1 : (\frac{M}{2}, n-1; \tau-1)$
- $C_2 : (\frac{M}{2}, n-1; \tau-1)$

<table>
<tr><td>0</td><td rowspan="6">$C^{'} : (M, n-1, \tau)$</td></tr>
<tr><td>$\vdots$</td></tr>
<tr><td>0</td></tr>
<tr><td>1</td></tr>
<tr><td>$\vdots$</td></tr>
<tr><td>1</td></tr>
</table>

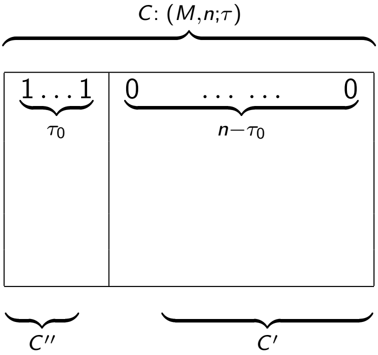
<table>
<tr><td>0</td><td rowspan="3">$C_1 : (\frac{M}{2}, n-1, \tau-1)$</td></tr>
<tr><td>$\vdots$</td></tr>
<tr><td>0</td></tr>
<tr><td>1</td><td rowspan="3">$C_2 : (\frac{M}{2}, n-1, \tau-1)$</td></tr>
<tr><td>$\vdots$</td></tr>
<tr><td>1</td></tr>
</table>

- Пресмятаме тегловите спектри от система (3):
  - $w = (w_0, w_1, \dots, w_n)$ на $C : OA(M,n;\tau)$
  - $w^{'} = (w^{'}_0, w^{'}_1, \dots, w^{'}_{n-1})$ на $C^{'} : OA(M,n-1;\tau)$
  - $w^1 = (w^1_0, w^1_1, \dots, w^1_{n-1})$ на $C_1 : OA(\frac{M}{2},n-1;\tau-1)$
  - $w^2 = (w^2_0, w^2_1, \dots, w^2_{n-1})$ на $C_2 : OA(\frac{M}{2},n-1;\tau-1)$
- **основна Теорема.** За всеки теглови спектър w на C, могат да се определят свързаните с него теглови спектри $w^{'}$, w^1 и w^2 на $C^{'}$, C_1 и C_2.
- ограничения за броя на получените теглови спектри на C:$OA(M,n;\tau)$
- резултати за несъществуване
- нови конструкции

Дизайни в $H(n,2)$

Връзки между спектрите на дизайни с параметри $(M,n;\tau)$ и $(M,n-\tau_0;\tau)$

- C:$OA(M,n;\tau) \in H(n,2)$ дизайн
- τ_0: $\tau_0 \leq \tau, 2\tau_0 \leq n$, $\exists\, x \in C : wt(x) = \tau_0$
- C':$OA(M,n-\tau_0;\tau)$; C'':$OA(M,\tau_0;\tau_0)$;



- за фиксирано τ_0:$\tau_0 \leq \tau, 2\tau_0 \leq n$
- за всеки теглови спектър $w = (w_0, w_1, \ldots, w_n)$ на C
- тегловия спектър $w' = \{w'_1, w'_2, \ldots, w'_{n-\tau_0}\}$ на C':$OA(M, n-\tau_0; \tau)$;
- спектъра $p(x) = (p_0(x), \ldots, p_n(x))$ на точка $x \in C$ с тегло $wt(x) = \tau_0$
- числата $p_{i,j}(x) = |\{y \in C : wt(y) = j, d(x,y) = i\}$
- информация за структурата на i-блока на $C : OA(M, n; \tau)$
- ограничения за броя на получените теглови спектри на C:$OA(M, n; \tau)$

A. V. Khalyavin, *Estimates of the capacity of orthogonal arrays of large strength*, Moscow Univ. Math. Bull. 65, 2010, 130-131.

Таблица 2. Нови долни граници за $L(n, \tau)$

n \ τ	4	5	6	7	8	9	10
4	1						
5	1	1					
6	2	1	1				
7	${}^{sz}4$	2	1	1			
8	4^c	${}^{sz}4$	2	1	1		
9	6,7,8	4^c	4	2	1	1	
10	6,7,8	6,7,8	6,7,8	4	2	1	1
11	6,7,8	6,7,8	8^c	6,7,8	4	2	1
12	7,8	6,7,8	12–15,16	8^c	6,7,8	4	2
13	8	7,8	16	12–15,16	10–15,16	6,7,8	4
14	8	8	16	16	16^c	10–15,16	6,7,8

Peter Boyvalenkov et al., *Nonexistence of a few binary orthogonal arrays*, Discrete Applied Mathematics 217(2)

- за всяко $4 \le n \le 14$ и $\tau = 4, 5$

Таблица 3. Точни стойности за $L(n,\tau)$

n \ τ	4	5	6	7	8	9	10
4	1						
5	1	1					
6	2	1	1				
7	$^{sz}4$	2	1	1			
8	4^c	$^{sz}4$	2	1	1		
9	8	4^c	4	2	1	1	
10	8	8	8	4	2	1	1
11	8	8	8^c	8	4	2	1
12	8	8	12–15,16	8^c	8	4	2
13	8	8	16	12–15,16	16	8	4
14	8	8	16	16	16^c	16	8

БЛАГОДАРЯ ЗА ВНИМАНИЕТО!