

UNIVERZITA KOMENSKÉHO V BRATISLAVE

Fakulta matematiky, fyziky a informatiky



**POUŽITIE METÓD SEMIDEFINITNÉHO
PROGRAMOVANIA V OPTIMÁLNO
NAVRHOVANÍ EXPERIMENTOV**

Diplomová práca

Bratislava 2013

Bc. Erika Wagnerová

UNIVERZITA KOMENSKÉHO V BRATISLAVE

Fakulta matematiky, fyziky a informatiky



**POUŽITIE METÓD SEMIDEFINITNÉHO
PROGRAMOVANIA V OPTIMÁLNO
NAVRHOVANÍ EXPERIMENTOV**

Diplomová práca

Študijný program: Ekonomická a finančná matematika

Študijný odbor: 9.1.9 Aplikovaná matematika, 1114

Školiace pracovisko: Katedra aplikovanej matematiky a štatistiky

Vedúci diplomovej práce: Mgr. Lenka Filová, PhD.

Bratislava 2013

Bc. Erika Wagnerová



Univerzita Komenského v Bratislave
Fakulta matematiky, fyziky a informatiky


ZADANIE ZÁVEREČNEJ PRÁCE

Meno a priezvisko študenta: Bc. Erika Wagnerová
Študijný program: ekonomická a finančná matematika (Jednoodborové štúdium, magisterský II. st., denná forma)
Študijný odbor: 9.1.9. aplikovaná matematika
Typ záverečnej práce: diplomová
Jazyk záverečnej práce: slovenský

Názov: Použitie metód semidefinitného programovania v optimálnom navrhovaní experimentov

Cieľ: Aplikovať metódy semidefinitného programovania v tých situáciach v navrhovaní experimentov, kde kritérium optimality nie je diferencovateľné, a teda nie je možné použiť gradientné metódy. Postupy ilustrovať na konkrétnych modeloch

Vedúci: Mgr. Lenka Filová, PhD.
Katedra: FMFI.KAMŠ - Katedra aplikovanej matematiky a štatistiky
Vedúci katedry: prof. RNDr. Daniel Ševčovič, CSc.
Dátum zadania: 25.01.2012
Dátum schválenia: 26.01.2012


prof. RNDr. Daniel Ševčovič, CSc.
garant študijného programu



študent



vedúci práce

Čestné prehlásenie

Čestne prehlasujem, že predloženú diplomovú prácu som vypracovala samostatne pod odborným vedením vedúcej diplomovej práce a s použitím uvedenej literatúry.

V Bratislave, apríl 2013

.....

Erika Wagnerová

Podakovanie

Týmto chcem poďakovať vedúcej diplomovej práce Mgr. Lenke Filovej, PhD. za jej odborné vedenie, cenné rady, pripomienky, podporu a množstvo času, ktorý mi venovala počas písania tejto práce.

Abstrakt

Wagnerová, Erika: **Použitie metód semidefinitného programovania v optimálnom navrhovaní experimentov**. [Diplomová práca]. Univerzita Komenského v Bratislave. Fakulta matematiky, fyziky a informatiky; Katedra aplikovanej matematiky a štatistiky.

Vedúci diplomovej práce: Mgr. Lenka Filová, Phd.

Bratislava: FMFI UK, 2013. /43 s./

Cieľom tejto práce je výpočet aproximatívneho návrhu (A-optimálnych a V-optimálnych návrhov) experimentu s lineárnym modelom pomocou semidefinitného programovania. Je známe, že hľadanie A-optimálneho návrhu možno formulovať ako problém semidefinitného programovania, teda je možné riešiť ho v polynomickej čase. Kľúčovou myšlienkou je ukázať, že po vhodnej parametrizácii je V-optimálny návrh ekvivalentný A-optimálnemu. Výpočet V-optimálneho návrhu je ekvivalentný výpočtu A-optimálneho návrhu pre parametrizovaný model. Formulácia pomocou semidefinitného programovania je výhodná, pretože umožňuje do úlohy zahrnúť dodatočné konvexné ohraničenia na návrh, ako napríklad marginálne ohraničenia a ohraničenia na cenu. Problém výpočtu V-optimálneho návrhu pomocou semidefinitného programovania ilustrujeme na príkladoch ako sú polynomickej model, central composite design a kvadratický model s marginálnymi ohraničeniami a ohraničeniami na cenu.

Kľúčové slová: semidefinitné programovanie, A-optimálny návrh, V-optimálny návrh, reparametrizácia modelu, marginálne ohraničenia, ohraničenia na cenu.

Abstract

Wagnerová, Erika: **Using semidefinite programming methods in optimal design of experiments**. [Master's thesis]. Comenius University Bratislava. Faculty of mathematics, physics and informatics; Department of applied mathematics and statistics.

Supervisor: Mgr. Lenka Filová, Phd.

Bratislava: FMFI UK, 2013. /43 p./

The aim of this work is to compute approximate experimental design (A-optimal and V-optimal design) for linear regression models by semidefinite programming. It is well known that, the A-optimal design problem can be cast as semidefinite programming problem. Thus it can be solved in polynomial time. The key idea is to show that V-optimality is equivalent to A-optimality under a particular parametrization. Computing V-optimal design is equivalent to computing A-optimal design for parametrization of the model. The formulation of optimal design as semidefinite programming problem has advantage that we can easily include convex constraints on the design, for instance the marginal and cost constraints. Problem of computing V-optimal design by semidefinite programming is illustrated with examples such as polynomial model, central composite design and quadratic model with marginal and cost constraints.

Keywords: semidefinite programming, A-optimal design, V-optimal design, parametrization of model, marginal constraints, cost constraints.

Obsah

Úvod	9
1 Optimálne navrhovanie experimentov	10
1.1 Lineárny model	10
1.2 Návrh experimentov	11
1.3 Kritériá optimality	13
2 Semidefinitné programovanie	16
2.1 Zovšeobecnenie úlohy lineárneho programovania	16
2.2 Formulácia úlohy SDP a jej duálnej úlohy	16
2.3 Dualita v SDP	18
2.4 Formulácia kritérií optimality pomocou konvexného programovania	20
3 A-optimality vs. V-optimality	23
3.1 V-optimality	23
3.2 Formulácia V-optimality ako L-optimality	24
3.3 Parametrizácia lineárneho modelu (1.1)	24
3.4 Eficiencia A-optimálneho a V-optimálneho návrhu	26
3.5 Formulácia V-optimality ako SDP	26
3.6 Parametrizácia response surface modelu	27
4 Príklady a aplikácie	28
4.1 Legendrova parametrizácia pre kvadratický model na intervale $[-1, 1]$	28
4.2 Legendrova parametrizácia pre kvadratický model na štvorci $[-1, 1]^2$	29
4.3 V-optimálny návrh pre polynomický model stupňa d na intervale $[-1, 1]$	33
4.4 Proces zhusťovania uránového paliva	36
4.5 Central Composite Design	42
Záver	47
Literatúra	48
Príloha	50

Úvod

Semidefinitné programovanie (SDP) patrí do triedy úloh konvexného programovania a zjednocuje niekoľko štandardných optimalizačných problémov, ako napríklad lineárne a kvadratické konvexné programovanie. Jedná sa o minimalizáciu lineárnej funkcie vzhľadom na lineárnu kombináciu symetrických matíc, ktorá spĺňa nelineárnu podmienku kladnej semidefinitnosti. SDP bolo počas 90. rokov 20. storočia jednou z najaktívnejších oblastí výskumu optimalizácie. Dôvodom tejto popularity bol hlavne rozvoj efektívnych algoritmov, metód vnútorného bodu, ktoré konvergujú k optimálnemu riešeniu v polynomicom čase v kombinácii s množstvom aplikácií, kde SDP možno použiť na riešenie NP-ťažkých problémov. SDP nájde využitie v oblastiach ako kombinatorická optimalizácia, optimálne riadenie a v neposlednom rade optimálny návrh experimentov. Je známe, že napríklad kritériá A- a E-optimality je možné formulovať ako úlohu SDP. Takáto formulácia navyše umožňuje zahrnúť do úlohy dodatočné konvexné ohraničenia na návrh, napríklad na celkovú cenu experimentu.

Cieľom tejto práce je aplikovať metódy semidefinitného programovania v situáciách, kde zlyhávajú gradientné metódy výpočtu optimálneho návrhu (napr. V-optimálny návrh, úloha s dodatočnými ohraničeniami na návrh). Následne tieto postupy ilustrovať na konkrétnych modeloch, pričom sa zaoberáme len lineárnymi regresnými modelmi.

Práca pozostáva zo štyroch kapitol. V prvej kapitole sa zaoberáme teóriou optimálneho navrhovania experimentov. Druhá kapitola je venovaná základnej teórii semidefinitného programovania a formulácii kritérií optimality pomocou konvexného programovania. Tretiu kapitolu venujeme vzťahu A-optimality a V-optimality. Ukazujeme, za akých podmienok možno kritérium V-optimality previesť na kritérium A-optimality, a teda riešiť ako úlohu SDP. Inými slovami po vhodnej parametrizácii pôvodného modelu vieme kritérium V-optimality formulovať ako kritérium A-optimality v novom modeli. Štvrtá kapitola pozostáva z príkladov, kde cieľom je nájsť V-optimálny návrh pre modely ako sú polynomický model stupňa d na intervale, kvadratický response surface model s dodatočnými ohraničeniami na celkovú cenu experimentu a marginálnymi ohraničeniami a central composite design.

1 Optimálne navrhovanie experimentov

V tejto kapitole sa oboznámime so základnou teóriou optimálneho navrhovania experimentov. Definujeme si lineárny model a uvedieme niekoľko predpokladov, ktoré sa s týmto modelom spájajú. Vysvetlíme rozdiel medzi exaktným a aproximatívnym návrhom a uvedieme ich zápis. Uvedieme niekoľko kritérií optimality, ktorým sa budeme venovať aj v ďalších kapitolách. Teóriu sme čerpali z [2] a [11].

1.1 Lineárny model

Majme experiment, ktorý pozostáva z N meraní. Naším cieľom je čo najpresnejšie odhadnúť vektor neznámych parametrov β . V bodoch x_1, x_2, \dots, x_n uskutočníme r_1, r_2, \dots, r_n meraní premennej y , pričom $\sum_{i=1}^n r_i = N$. V týchto bodoch nameriame hodnoty $y(x_i)$:

$$y(x_i) = f(x_i)^T \beta + \varepsilon_i, \quad i = 1, \dots, n, \quad (1.1)$$

kde $y(x_i)$ je pozorovanie v bode x_i , β je m -rozmerný vektor neznámych parametrov, $f(x_i)^T = (f_1(x_i), f_2(x_i), \dots, f_m(x_i))^T$ sú známe funkcie, ktoré sú spojité na \mathcal{X} . ε_i je náhodná chyba v i -tom meraní. Predpokladáme, že chyby jednotlivých meraní sú nekorelované t.j.:

$$E(\varepsilon_i) = 0 \quad \text{pre } i = 1, \dots, n \quad \text{a} \quad \text{cov}(\varepsilon_i, \varepsilon_j) = \begin{cases} 0 & \text{ak } i = j \\ \sigma^2 & \text{ak } i \neq j \end{cases}.$$

Lineárny model (1.1) je možné zapísať maticovo:

$$\mathbf{y} = \mathbf{F}\beta + \varepsilon, \quad (1.2)$$

kde

$$\mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}, \quad \mathbf{F} = \begin{pmatrix} f_1(x_1) & f_2(x_1) & \dots & f_m(x_1) \\ f_1(x_2) & f_2(x_2) & \dots & f_m(x_2) \\ \vdots & \vdots & \vdots & \vdots \\ f_1(x_n) & f_2(x_n) & \dots & f_m(x_n) \end{pmatrix} = \begin{pmatrix} f^T(x_1) \\ f^T(x_2) \\ \vdots \\ f^T(x_n) \end{pmatrix},$$
$$\beta = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_m \end{pmatrix}, \quad \varepsilon = \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix}.$$

Ako sme už spomínali, β je vektor neznámych parametrov, ktorý možno odhadnúť pomocou metódy najmenších štvorcov. Odhad $\hat{\beta}$ je teda riešením optimalizačnej úlohy:

$$\min_{\beta \in \mathbb{R}^m} \|\mathbf{y} - \mathbf{F}\beta\| = \min_{\beta \in \mathbb{R}^m} (\mathbf{y} - \mathbf{F}\beta)^T (\mathbf{y} - \mathbf{F}\beta) = \min_{\beta \in \mathbb{R}^m} \mathbf{y}^T \mathbf{y} - 2\mathbf{y}^T \mathbf{F}\beta + (\mathbf{F}\beta)^T (\mathbf{F}\beta) \quad (1.3)$$

Derivovaním (1.3) dostaneme systém normálnych rovníc:

$$\mathbf{F}^T \mathbf{F}\beta = \mathbf{F}^T \mathbf{y}$$

Za predpokladu, že \mathbf{F} s rozmermi $n \times m$ je nenáhodná matica s plnou hodnotou t.j. $h(\mathbf{F}) = m$, odhad vektora parametrov bude mať tvar:

$$\hat{\beta} = (\mathbf{F}^T \mathbf{F})^{-1} \mathbf{F}^T \mathbf{y}$$

O strednej hodnote a variancii odhadu $\hat{\beta}$ platí:

$$E(\hat{\beta}) = \beta, \quad Var(\hat{\beta}) = \sigma^2 (\mathbf{F}^T \mathbf{F})^{-1},$$

pričom ak našim cieľom je porovnávanie niekoľkých návrhov jedného experimentu, hodnota σ^2 nie je podstatná, pretože táto hodnota bude pre všetky návrhy konkrétneho experimentu rovnaká. Pre podrobnejšie vysvetlenie problematiky pozri [11].

1.2 Návrh experimentov

Majme model $y(x) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \varepsilon$ a chceme nájsť D-optimálny návrh (pozri 1.3.3) na množine $\mathcal{X} = [-1, 1]$, pričom N je počet meraní. D-optimálny návrh určí, že polovicu meraní máme uskutočniť v bode $x = -1$ a polovicu meraní v bode $x = 1$. Pre N párne to znamená, že v bode $x = -1$ aj v bode $x = 1$ uskutočnime $\frac{N}{2}$ meraní. Pokiaľ N nie je párne, situácia sa komplikuje a výsledný návrh bude závisieť od hodnoty N .

Predpokladáme, že $x_1, x_2, \dots, x_n \in \mathcal{X}$, kde \mathcal{X} označuje výberový priestor experimentu, pričom \mathcal{X} je kompaktná množina. Ξ označuje množinu všetkých návrhov na \mathcal{X} .

Definícia 1.1. *Exaktný návrh pre N meraní v n rôznych bodoch definujeme ako mieru*

$$\xi_N = \left\{ \begin{array}{cccc} x_1 & x_2 & \dots & x_n \\ r_1/N & r_2/N & \dots & r_n/N \end{array} \right\},$$

kde r_i sú celé čísla počtov meraní v bodoch x_i a platí $\sum_{i=1}^n r_i = N$.

Definícia 1.2. *Informačnú maticu exaktného návrhu pre N meraní definujeme ako*

$$M(\xi_N) = \frac{F^T F}{N} = \frac{\sum_{i=1}^n r_i f(x_i) f(x_i)^T}{N}.$$

F je matica z lineárneho modelu (1.2).

V praxi potrebujeme pracovať s exaktnými návrhmi. Problém nájdenia optimálneho návrhu experimentu sa zjednodušuje na hľadanie aproximatívneho návrhu, čím zanedbáme podmienku celočíselnosti počtov meraní v bodoch x_i . Potom je možné nájsť z aproximatívneho návrhu exaktný za pomoci celočíselnej aproximácie (napr. Pukelsheimova zaokrúhľovacia metóda pozri [12]).

Definícia 1.3. *Aproximatívny návrh je vyjadrený mierou ξ na \mathcal{X} , ktorú v n rôznych bodoch definujeme ako*

$$\xi = \left\{ \begin{array}{cccc} x_1 & x_2 & \dots & x_n \\ w_1 & w_2 & \dots & w_n \end{array} \right\},$$

kde $\xi(x_i) = w_i$ sú váhy v bodoch x_i , pre ktoré platí

$$0 \leq w_i \leq 1 \quad i = 1, \dots, n$$

$$\sum_{i=1}^n w_i = 1.$$

ξ je pravdepodobnostná miera, preto musí pre ňu platiť podmienka $\int_{\mathcal{X}} \xi(dx) = 1$.

Definícia 1.4. *Informačnú maticu aproximatívneho návrhu definujeme ako*

$$M(\xi) = \int_{\mathcal{X}} f(x)f(x)^T \xi(dx) = \sum_{i=1}^n w_i f(x_i)f(x_i)^T.$$

Varianciu odhadu odpovede y možno zapísať nasledovne

$$Var\{\hat{y}(x)\} = Var\{f^T(x)\hat{\beta}\} = f^T(x)Var\{\hat{\beta}\}f(x) = \sigma^2 f^T(x)(F^T F)^{-1}f(x).$$

Pri porovnávaní rôznych návrhov experimentu je často užitočné varianciu preškálovať vzhľadom na σ^2 a vziať do úvahy počet opakovaní experimentu N , čím dostávame variančnú funkciu

$$d(x, \xi) = N \frac{Var\{\hat{y}(x)\}}{\sigma^2}.$$

Definícia 1.5. *Variančná funkcia exaktného návrhu je definovaná ako*

$$d(x, \xi_N) = f(x)^T M^{-1}(\xi_N)f(x).$$

Definícia 1.6. *Variančná funkcia aproximatívneho návrhu je definovaná ako*

$$d(x, \xi) = f(x)^T M^{-1}(\xi)f(x).$$

1.3 Kritériá optimality

Na základe kritérií optimality sa rozhodujeme, ktorý návrh je pre nás dostatočne dobrý. Teória ponúka širokú škálu kritérií optimality. Mnohé z nich sú nazvané podľa písmen abecedy. My upriamime pozornosť len na niekoľko z mnoha kritérií. Viac napríklad v [2].

Definícia 1.7. *Kritériom optimality nazývame funkciu*

$$\Psi : \{\mathcal{M}\} \rightarrow \langle 0, \infty \rangle,$$

kde \mathcal{M} je označuje množinu všetkých informačných matíc.

Návrh ξ^* je Ψ -optimálny práve vtedy, keď $\xi^* = \arg \max_{\xi \in \Xi} \Psi\{M(\xi)\}$.

$M(\xi^*)$ nazývame Ψ -optimálnou informačnou maticou.

Symbolmi S_{++}^m a S_+^m označme množinu symetrických kladne definitných resp. kladne semidefinitných matíc typu $m \times m$.

Definícia 1.8. *Kritérium Ψ je pozitívne homogénne, ak spĺňa podmienku*

$$\Psi(\alpha C) = \alpha \Psi(C) \quad \forall C \in S_+^m, \alpha \geq 0.$$

Vlastnosť pozitívnej homogenity potrebujeme, aby sme mohli korektne definovať efícienciu návrhu, ktorú budeme potrebovať na meranie kvality návrhu.

Definícia 1.9. *Efícienciu návrhu ξ vzhľadom na kritérium Ψ definujeme nasledovne*

$$eff(M(\xi)|\Psi) = \frac{\Psi(M(\xi))}{\max_{\zeta \in \Xi} \Psi(M(\zeta))}.$$

Uvedieme si niekoľko najznámejších kritérií optimality, ktoré sú v praxi najčastejšie používané (D- prípadne A-optimality) a doplníme ich o ďalšie kritériá, ktoré sa dajú formulovať ako úlohy konvexného programovania (LP, SDP, maxdet).

1.3.1 A-optimality („average“)

Kritérium A-optimality je definované ako

$$\Phi_A(M(\xi)) = \begin{cases} 0 & \text{ak } M \in S_+^m \setminus S_{++}^m \\ m(tr(M(\xi)^{-1}))^{-1} & \text{ak } M \in S_{++}^m. \end{cases}$$

A-optimálny návrh minimalizuje súčet variancií odhadov parametrov

$$\Phi_A(M(\xi^*)) = \min \sum_{i=1}^m \lambda_i^{-1},$$

kde λ_i $i = 1, \dots, m$ sú vlastné hodnoty informačnej matice.

1.3.2 c-optimalita

Zaujímá nás odhad lineárnej kombinácie parametrov modelu t.j. $c^T \beta$ s minimálnou variáciou, kde $c \in R^m$ je známy vektor konštánt a β je vektor neznámych parametrov. Keďže maximalizácia kritéria by v tomto prípade nemala vhodnú interpretáciu volíme prístup minimalizácie a optimálny návrh ξ^* hľadáme ako argument minima. Kritérium c-optimality je definované ako

$$\Psi_c(M(\xi)) = \begin{cases} \infty & \text{ak } c \text{ nepatrí do stĺpcového priestoru } M \\ \text{var } c^T \hat{\beta} \propto c^T M(\xi)^{-1} c & \text{ak } c \text{ patrí do stĺpcového priestoru } M. \end{cases}$$

Kritérium možno formulovať ako úlohu LP [7].

1.3.3 D-optimalita („determinant“)

Toto kritérium je jedno z najdôležitejších a najčastejšie používaných v praxi. D-optimálny návrh ξ^* maximalizuje

$$\Phi_D(M(\xi)) = \begin{cases} 0 & \text{ak } M \in S_+^m \setminus S_{++}^m \\ \det(M(\xi))^{1/m} & \text{ak } M \in S_{++}^m. \end{cases}$$

Prípadne možno použiť konvexnú verziu kritéria, kde sa minimalizuje

$$\Phi_D(M(\xi)) = -\ln(\det M(\xi)).$$

D-optimálny návrh minimalizuje zovšeobecnenú varianciu odhadov parametrov

$$\Phi_D(M(\xi^*)) = \min \prod_{i=1}^m \lambda_i^{-1},$$

kde λ_i sú vlastné hodnoty informačnej matice. D-optimálny návrh je invariantný vzhľadom na nedegenerované lineárne transformácie t.j. D-optimálny návrh modelu $y(x) = \beta^T f(x)$ je totožný s D-optimálnym návrhom modelu $y(x) = (\beta A)^T f(x)$, kde $\det A \neq 0$.

1.3.4 E-optimalita („eigenvalue“ alebo „extreme“)

E-optimálny návrh maximalizuje najmenšiu vlastnú hodnotu informačnej matice t.j. maximalizuje

$$\Phi_E(M(\xi)) = \begin{cases} 0 & \text{ak } M \in S_+^m \setminus S_{++}^m \\ \lambda_{\min}(M(\xi)) & \text{ak } M \in S_{++}^m. \end{cases}$$

1.3.5 V-optimalita („variance“)

Kritérium minimalizuje priemernú varianciu na oblasti \mathcal{R} , pričom priemerovanie robíme vzhľadom na pravdepodobnostné rozdelenie μ na oblasti \mathcal{R} . Opäť kvôli vhodnejšej interpretácii optimálny návrh ξ^* budeme hľadať ako argument minima. Kritérium bude vyzerat

$$\Phi_V(M(\xi)) = \int_{\mathcal{R}} f^T(x)M^{-1}(\xi)f(x)\mu(dx) = \int_{\mathcal{R}} d(x, \xi)\mu(dx). \quad (1.4)$$

Tomuto kritériu sa budeme podrobnejšie venovať v kapitole 3.

2 Semidefinitné programovanie

V tejto kapitole sa budeme venovať základom problematiky semidefinitného programovania (alebo SDP). Sformulujeme štandardnú úlohu SDP a jej duálnu úlohu, vysvetlíme základné princípy duality v SDP a naznačíme niekoľko aplikácií SDP. Pre viac informácií pozri napríklad [1], [3], [4] alebo [13].

Záujem o štúdium SDP vzrástol hlavne vďaka objaveniu nových aplikácií v rôznych oblastiach ako napríklad kombinatorická optimalizácia (MAX CUT problém), teória optimálneho riadenia, štatistika, optimálne navrhovanie experimentov (viac v [3] alebo [13]) v kombinácii s rozvojom nových efektívnych algoritmov pre riešenie úloh SDP pomocou metód vnútorného bodu, ktoré k optimálnemu riešeniu konvergujú v polynomickej čase.

2.1 Zovšeobecnenie úlohy lineárneho programovania

[13] uvádza, že na úlohu SDP sa dá pozrieť ako na rozšírenie úlohy lineárneho programovania (LP)

$$\begin{aligned} \min_x \quad & c^T x \\ & a_i^T x = b_i, \quad i = 1, \dots, m \\ & x \geq 0 \end{aligned}$$

Podmienka nezápornosti $x \geq 0$ je v prípade SDP nahradená zovšeobecnenou nerovnosťou vzhľadom na priestor kladne semidefinitných matic. Vďaka tejto nelineárnej podmienke je SDP oveľa všeobecnejšie ako LP a širokú škálu konvexných úloh možno formulovať a veľmi efektívne riešiť ako úlohu SDP.

2.2 Formulácia úlohy SDP a jej duálnej úlohy

V úlohe SDP optimalizujeme lineárnu funkciu symetrickej matice X , vzhľadom na m lineárnych maticových nerovností za podmienky kladnej semidefinitnosti matice X . Takáto podmienka je síce nelineárna a nehladká, ale konvexná.

Problém možno zapísať v tvare

$$\begin{aligned} \min \quad & C \bullet X \\ & A_i \bullet X = b_i, \quad i = 1, \dots, m \\ & X \succeq 0 \end{aligned} \tag{2.1}$$

$X \in S^n$, čo je priestor symetrických $n \times n$ matíc, $b \in R^m$ je známy vektor, matice $A_i \in S^n$ a $C \in S^n$ sú tiež známe.

Nerovnosť $X \succeq 0$ značí obmedzenie matice X na kladne semidefinitnú maticu, ohraničenie " \succeq " indikuje Löwnerove parciálne usporiadanie t.j. pre reálne symetrické matice A a B platí že, $A \succ B$ ($A \succeq B$), keď matica $A - B$ je kladne definitná (kladne semidefinitná). Zápis $C \bullet X$ označuje skalárny súčin symetrických matíc C a X

$$C \bullet X = \text{tr}(C^T X) = \sum_{i,j} C_{ij} X_{ij}.$$

Všimnime si, že ak matica X by bola diagonálna, úloha (2.1) sa redukuje na úlohu LP.

Bez ujmy na všeobecnosti predpokladáme v úlohe (2.1) symetrickosť matíc C a A_i . Ak by C nebola symetrická, bolo by možné maticu C v (2.1) nahradiť maticou $\frac{1}{2}(C + C^T)$, ktorá už symetrická je. Keďže

$$C \bullet X = \text{tr}(C^T X) = \text{tr}((C^T X)^T) = \text{tr}(XC) = \text{tr}(CX) = C^T \bullet X,$$

tak

$$\frac{1}{2}(C + C^T) \bullet X = \frac{C \bullet X + C^T \bullet X}{2} = C \bullet X.$$

Obdobná argumentácia platí aj pre matice A_i .

Predpoklad symetrickosti nám umožňuje formulovať **duálnu úlohu k** (2.1) v nasledovnom tvare

$$\begin{aligned} \max \quad & \sum_{i=1}^m b_i y_i \\ & \sum_{i=1}^m A_i y_i + S = C \\ & S \succeq 0 \end{aligned} \tag{2.2}$$

premené $y \in R^m$ a $S \in S^n$. Bez ujmy na všeobecnosti predpokladáme, že A_i pre $i = 1, \dots, m$ sú lineárne nezávislé. Všimnime si, že $S = C - \sum_{i=1}^m A_i y_i$, preto úlohu (2.2) možno zapísať aj v tvare

$$\begin{aligned} \max \quad & \sum_{i=1}^m b_i y_i \\ & C - \sum_{i=1}^m A_i y_i \succeq 0 \end{aligned} \tag{2.3}$$

Podobne ako úloha LP, tak aj úlohu SDP možno zapísať v rôznych formách. My sme uvieli jej štandardnú formu, ktorá je jednou z mnohých. Niekedy môže byť podmienka kladnej semidefinitnosti aplikovaná na lineárnu kombináciu viacerých matíc, ako napríklad v (2.3), inokedy sa môže v úlohe vyskytnúť podmienka nezápornosti pre skalárne premenné alebo Löwnerova nerovnosť pre maticové premenné. Môžeme mať obmedzenia na kladnú semidefinitnosť na niekoľko maticových výrazov, prípadne kombináciu niekoľkých alebo všetkých uvedených ohraničení naraz. Takúto úlohu je možné potom previesť na štandardnú úlohu SDP, obyčajne zavedením nových premenných, či už skalárnych alebo maticových, prípadne zavedením nových ohraničení. Viac o formulácii primárnej a duálnej úlohy v SDP v [1] a [13].

2.3 Dualita v SDP

Na začiatok tejto podkapitoly si zavedieme niekoľko dôležitých označení, ktoré budeme neskôr potrebovať:

- p^* resp. d^* sú optimálne hodnoty úloh (2.1) resp. (2.3).
- Úloha (2.1) je prípustná, ak množina $\{X \in S^m | A_i \bullet X = b_i \quad i = 1, \dots, m, X \succeq 0\}$ je neprázdna.
- V opačnom prípade je (2.1) neprípustná.
- Úloha (2.3) je prípustná, ak množina $\{y \in R^m | C - \sum_{i=1}^m A_i y_i \succeq 0\}$ je neprázdna.
- V opačnom prípade je (2.3) neprípustná.
- Infimum na práznej množine definujeme ako $+\infty$ t.j. ak je úloha (2.1) neprípustná, tak $p^* = +\infty$.
- Supremum na prázdnej množine definujeme ako $-\infty$ t.j. ak je úloha (2.3) neprípustná, tak $d^* = -\infty$.
- Primárna úloha (2.1) je neohraničená, ak $\inf\{C \bullet X | A_i \bullet X = b_i \quad i = 1, \dots, m, X \succeq 0\}$ je rovné $-\infty$.
- Duálna úloha (2.3) je neohraničená, ak $\sup\{b^T y | C - \sum_{i=1}^m A_i y_i \succeq 0\}$ je rovné $+\infty$.

Pre zjednodušenie formulácie úloh SDP zavedieme nasledovné značenie:

Ak A je $p \times q$ matica, potom $\text{vec}A$ je vektor dĺžky pq , vytvorený zo stĺpcov matice A

nasledovne

1. Nech $A = \left(\begin{array}{c|c|c|c} & & \dots & \\ \hline a_1 & a_2 & \dots & a_q \\ \hline & & \dots & \end{array} \right)$, kde a_1, a_2, \dots, a_q sú stĺpce matice A dĺžky p , potom

$$\text{vec}A = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_q \end{pmatrix}.$$

2. Ak v je vektor dĺžky pq , potom $\text{Mat}_{pq}v$ je $p \times q$ matica, ktorej i -ty stĺpec je vytvorený z $(i-1)p+1$ -teho až ip -teho prvku vektora v .

Napríklad sústavu rovností

$$A_i \bullet X = b_i \quad \text{pre } i = 1, \dots, m \quad \text{možno jednoducho zapísať ako}$$

$$\mathcal{A} \text{vec}X = \mathbf{b},$$

kde \mathcal{A} je matica typu $m \times n^2$ t.j.

$$\mathcal{A} = \begin{pmatrix} \text{vec}^T(A_1) \\ \text{vec}^T(A_2) \\ \dots \\ \text{vec}^T(A_m) \end{pmatrix} \quad \text{a } \mathbf{b} = \begin{pmatrix} b_1 \\ b_2 \\ \dots \\ b_m \end{pmatrix}.$$

Vďaka polyédrickému charakteru prípustnej množiny, teória duality v LP predstavuje veľmi silný výsledok. Optimálne hodnoty primárnej a duálnej úlohy v LP sú vždy rovnaké, až na patologický prípad, keď obe úlohy sú neprípustné. Treba si uvedomiť, že výsledky duality v prípade SDP sú slabšie, čo uvidíme nižšie. Slabá veta o dualite a silná veta o dualite pre SDP je podobné ako v LP. Uvedieme ich znenie.

Veta 2.1. *Nech X je ľubovoľná prípustná matica primárnej úlohy (2.1) a nech vektor y je ľubovoľný prípustný vektor duálnej úlohy (2.3). Potom $C \bullet X \geq b^T y$.*

Veta 2.2. *Nech p^* resp. d^* označujú optimálne hodnoty úloh (2.1) resp. (2.3). Potom $p^* = d^*$, ak platia nasledujúce podmienky:*

1. *Ak primárna úloha (2.1) je striktne prípustná, t.j. množina $\{X \in S^n \mid \mathcal{A} \text{vec}X = \mathbf{b} \quad i = 1, \dots, m, X \succ 0\}$ je neprázdna.*

2. Ak duálna úloha (2.3) je striktne prípustná t.j. množina $\{y \in R^m | C - \text{Mat}(\mathcal{A}^T \mathbf{y}) \succ 0\}$ je neprázdna.

Je dôležité uvedomiť si úskalia teórie duality v SDP. Z vety o dualite 2.2 vyplýva, že pri porušení ľubovoľnej podmienky nemusí platiť $p^* = d^*$, ktorá v LP platí. V kontraste s teóriou duality v LP je aj fakt, že v SDP možno dosiahnuť nenulovú duálnu medzeru. Duálnu medzeru v SDP možno zapísať ako

$$p - d = C \bullet X - b^T y = C \bullet X - \sum_{i=1}^m y_i A_i \bullet X = (C - \sum_{i=1}^m y_i A_i) \bullet X = S \bullet X.$$

Pokiaľ platí slabá veta o dualite 2.1 a primárna aj duálna úloha sú striktne prípusté, potom duálna medzera $S \bullet X = 0$.

V súvislosti s dualitou ešte spomenieme vetu o komplementarite, ktorej znenie je analogické so znením vety o komplementarite v LP.

Veta 2.3. *Nech X^* je prípustné riešenie primárnej úlohy (2.1) a y^* je prípustné riešenie duálnej úlohy (2.3). Nech $S^* = C - \text{Mat}(\mathcal{A}^T \mathbf{y}^*)$. Potom X^* a y^* sú optimálne riešenia primárnej a duálnej úlohy práve vtedy, keď $X^* S^* = 0$.*

Pre hlbšie pochopenie teórie semidefinitného programovania pozri [1], [13] a [4].

2.4 Formulácia kritérií optimality pomocou konvexného programovania

Niektoré z problémov optimálneho navrhovania experimentov môžu byť formulované a efektívne riešené ako úloha konvexného programovania, ktoré zahŕňajú aj úlohy SDP a maxdet. Takáto formulácia je veľmi výhodná vtedy, keď kritérium optimality nie je diferencovateľné (napr. E-optimality), a teda nie je možné použiť gradientné metódy. Formulácie úloh optimálneho navrhovania experimentov pomocou SDP či maxdet navyše umožňujú veľmi jednoducho do úlohy zahrnúť dodatočné konvexné ohraničenia na návrh. Viac o problematike v [3].

Predpokladajme diskretizáciu oblasti $\mathcal{X} = \{x_1, \dots, x_n\}$ a označme váhy $w_i = \xi(x_i)$. Návrh je teda vektor simplexu $\mathbb{S}_n = \{w \in R^n : w_1 + \dots + w_n = 1, w_1, \dots, w_n \geq 0\}$.

A-optimálny návrh

Maximalizácia stopy inverzie informačnej matice modelu sa dá formulovať ako úloha

SDP

$$\begin{aligned} \min \quad & \sum_{i=1}^m t_i \\ & \begin{pmatrix} \sum_{i=1}^n w_i f(x_i) f(x_i)^T & e_i \\ e_i^T & t_i \end{pmatrix} \succeq 0, \quad i = 1, \dots, m \\ & w \in \mathbb{S}_n, \end{aligned} \quad (2.4)$$

kde m je počet neznámych parametrov v modeli, $t_i \quad i = 1, \dots, m$ a $\lambda_i \quad i = 1, \dots, n$ sú neznáme a $e_i \in R^m$ je i -ty jednotkový vektor.

Prevod na úlohu SDP. Vychádzajme z úlohy hľadania A-optimálneho návrhu

$$\begin{aligned} \min \quad & \text{tr}(M^{-1}(\xi)) \\ & w \in \mathbb{S}_n. \end{aligned} \quad (2.5)$$

Diagonálne prvky matice $M^{-1}(\xi)$ možno zapísať v tvare

$$M_{ii}^{-1}(\xi) = e_i^T M^{-1}(\xi) e_i \quad i = 1, \dots, m.$$

Kritérium zapíšeme v tvare $\min \sum_{i=1}^m e_i^T M^{-1}(\xi) e_i$. Zavedením pomocných premenných t_1, \dots, t_m úloha (2.5) nadobudne tvar

$$\begin{aligned} \min \quad & \sum_{i=1}^m t_i \\ & e_i^T M^{-1}(\xi) e_i \leq t_i \quad i = 1, \dots, m \\ & w \in \mathbb{S}_n. \end{aligned}$$

Teda

$$\begin{aligned} \min \quad & \sum_{i=1}^m t_i \\ & t_i - e_i^T M^{-1}(\xi) e_i \geq 0 \quad i = 1, \dots, m \\ & w \in \mathbb{S}_n. \end{aligned}$$

Aplikujeme kritérium kladnej semidefinitnosti pre Schurov doplnok.

Veta 2.4. *Nech* $X = \begin{pmatrix} A & B \\ B^T & C \end{pmatrix}$.

Ak $A \succ 0$, *potom* $X \succeq 0 \iff$ *Schurov doplnok* $C - B^T A^{-1} B \succeq 0$.

Po tejto úprave problém hľadania A-optimálneho návrhu nadobúda tvar úlohy SDP

$$\begin{aligned} \min \quad & \sum_{i=1}^m t_i \\ & \begin{pmatrix} M(\xi) & e_i \\ e_i^T & t_i \end{pmatrix} \succeq 0, \quad i = 1, \dots, m \\ & w \in \mathbb{S}_n. \end{aligned}$$

D-optimálny návrh

Maximalizácia determinantu informačnej matice modelu sa dá formulovať ako úloha maxdet

$$\begin{aligned} \min \quad & \ln \det \left(\sum_{i=1}^n w_i f(x_i) f(x_i)^T \right)^{-1} \\ & w \in \mathbb{S}_n. \end{aligned}$$

E-optimálny návrh

E-optimálny návrh maximalizuje najmenšiu vlastnú hodnotu informačnej matice $M(\xi)$. Problém nájdenia E-optimálneho návrhu možno preformulovať do SDP nasledovne

$$\begin{aligned} \max \quad & t \\ & \sum_{i=1}^n w_i f(x_i) f(x_i)^T \succeq tI \\ & w \in \mathbb{S}_n. \end{aligned}$$

3 A-optimalita vs. V-optimalita

Kritérium A-optimality a V-optimality sme už definovali v podkapitole 1.3. V tejto kapitole sa budeme venovať podrobnejšie vzťahu medzi týmito dvoma kritériami a uvedieme spôsob akým možno kritérium V-optimality previesť na kritérium A-optimality, a teda problém riešiť pomocou SDP.

3.1 V-optimalita

Kritérium (1.4) minimalizuje priemernú varianciu odozvy vzhľadom na pravdepodobnostné rozdelenie μ na oblasti \mathcal{R} . Kritérium sa v literatúre uvádza pod názvom I-optimalita (kvôli „Integrated”) alebo V-optimalita (kvôli „Variance”). Numerická hodnota kritéria (1.4) sa obyčajne odhaduje pomocou priemerovania variancie na diskretizácii oblasti \mathcal{R} . Hľadanie V-optimálneho návrhu teda vedie na diskretizáciu kritéria (1.4) [2].

Keďže kritérium V-optimality minimalizuje priemernú varianciu odpovede vzhľadom na oblasť \mathcal{R} , jeho použitie má význam pri response surface návrhoch. Predmetom záujmu je práve odozva, ktorá je ovplyvňovaná p kvantitatívnymi faktormi. Odozva je spojitá kvantitatívna premenná a priemerná odozva (mean response) je hladká neznáma funkcia p faktorov. Tieto faktory nadobúdajú reálne hodnoty. Graficky je priemerná odozva, ako funkcia kombinácie p faktorov, je $p + 1$ rozmerná plocha, ktorá sa nazýva response surface. Viac o response surface návrhoch v [2] a [6].

V ďalších príkladoch a aplikáciách volíme pravdepodobnostné rozdelenie μ rovnomerné alebo normálne. Ak predpokladáme diskretizáciu \mathcal{R} v r bodoch, tak $\mu(x_i) = \frac{1}{r}$ pre rovnomerné rozdelenie a $\mu(x_i) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x_i^2}{2}}$ pre normálne rozdelenie.

Informačnú maticu možno zapísať v tvare

$$M(\xi) = \sum_{i=1}^n w_i f(x_i) f^T(x_i).$$

Diskretizácia kritéria (1.4) bude mať teda tvar

$$\min \int_{\mathcal{R}} f^T(x) M^{-1}(\xi) f(x) \mu(dx) \approx \min \sum_{i=1}^r \mu(x_i) f^T(x_i) \left(\sum_{j=1}^n w_j f(x_j) f^T(x_j) \right)^{-1} f(x_i).$$

Za dodatočných obmedzení na váhy optimalizačná úloha bude mať tvar

$$\min \sum_{i=1}^r \mu(x_i) f^T(x_i) \left(\sum_{j=1}^n w_j f(x_j) f^T(x_j) \right)^{-1} f(x_i) \tag{3.1}$$

$$w \in \mathbb{S}_n.$$

Úlohu (3.1) možno riešiť napríklad pomocou funkcie *fmincon* v programe MATLAB.

3.2 Formulácia V-optimality ako L-optimality

V tejto podkapitole uvedieme za akých predpokladov a ako možno kritérium V-optimality previesť na kritérium A-optimality.

Nech $L = \int_{\mathcal{R}} f(x)f(x)^T \mu(dx)$ je kladne semidefinitná matica, potom (1.4) je špeciálnym prípadom tzv. kritéria L-optimality

$$\Psi_L(M(\xi)) = \text{tr}\{M^{-1}(\xi)L\}, \quad (3.2)$$

kde L je $m \times q$ matica s hodnosťou $h(L) = s \leq q$, pozri [2].

Ak L je $m \times m$ matica s plnou hodnosťou, $h(L) = m$, potom L možno vyjadriť v tvare $L = AA^T$, kde A je $m \times m$ matica s hodnosťou m , a teda

$$\text{tr}\{M^{-1}(\xi)L\} = \text{tr}\{M^{-1}(\xi)AA^T\} = \text{tr}\{A^T M^{-1}(\xi)A\} \quad (3.3)$$

Podľa (3.3) kritérium V-optimality možno formulovať ako $\Psi_V(M(\xi)) = \text{tr}\{A^T M^{-1}(\xi)A\}$. Všimnime si, že pokiaľ by $A = I$, tak (3.3) sa redukuje na kritérium A-optimality.

3.3 Parametrizácia lineárneho modelu (1.1)

Formulácia kritéria V-optimality pomocou kritéria L-optimality (3.3) nás navádza na myšlienku parametrizácie pôvodného lineárneho modelu. Úloha hľadania V-optimálneho návrhu pre pôvodný lineárny model sa mení na hľadanie A-optimálneho návrhu pre parametrizovaný model, teda na problém SDP. Výsledný A-optimálny návrh pre parametrizovaný model bude zodpovedať V-optimálnemu návrhu pre pôvodný model.

Predpokladajme, že máme návrh v n rôznych bodoch s vektorom výsledných váh $w \in R^n$ t.j.

$$\xi = \begin{Bmatrix} x_1 & x_2 & \dots & x_n \\ w_1 & w_2 & \dots & w_n \end{Bmatrix}.$$

Majme lineárny model, pre ktorý chceme nájsť V-optimálny návrh

$$y(x_i) = f(x_i)^T \beta + \varepsilon_i, \quad i = 1, \dots, n. \quad (3.4)$$

Nech parametrizovaný model, pre ktorý chceme nájsť A-optimálny návrh má tvar

$$y(x_i) = g(x_i)^T \gamma + \varepsilon_i, \quad i = 1, \dots, n. \quad (3.5)$$

Informačné matice prislúchajúce daným modelom budeme označovať

$$M_f(\xi) = \sum_{i=1}^n w_i f(x_i) f(x_i)^T \quad \text{a} \quad M_g(\xi) = \sum_{i=1}^n w_i g(x_i) g(x_i)^T.$$

Kritériá optimality prislúchajúce daným modelom vyzerajú nasledovne

$$\Psi_V(M_f(\xi)) = \text{tr}\{A^T M_f^{-1}(\xi) A\} \quad \text{a} \quad \Psi_A(M_g(\xi)) = \text{tr}\{M_g^{-1}(\xi)\}.$$

Potrebuje nájst vhodnú parametrizáciu, aby výsledný V-optimálny a A-optimálny návrh pre model (3.4) a (3.5) boli zhodné, teda chceme

$$\begin{aligned} \Psi_V(M_f(\xi)) &= \Psi_A(M_g(\xi)) \\ \text{tr}\{A^T M_f^{-1}(\xi) A\} &= \text{tr}\{M_g^{-1}(\xi)\} \\ A^T M_f^{-1}(\xi) A &= M_g^{-1}(\xi) \end{aligned}$$

Vhodnú parametrizáciu nám pomôže nájst nasledujúca veta.

Veta 3.1. *Nech $f(x)$ a $g(x)$ sú vektory regresných funkcií v modeloch (3.4) a (3.5). Nech $L = \int_{\mathcal{R}} f(x)f(x)^T \mu(dx)$, $L = AA^T$ a A je $m \times m$ matica s plnou hodnotou. Potom ak $g(x) = A^{-1}f(x)$, tak $M_g^{-1}(\xi) = A^T M_f^{-1}(\xi) A$.*

Dôkaz. $M_f(\xi) = \sum_{i=1}^n f(x_i)f(x_i)^T \xi(x_i)$

$$\begin{aligned} M_g(\xi) &= \sum_{i=1}^n g(x_i)g(x_i)^T \xi(x_i) \\ &= \sum_{i=1}^n A^{-1}f(x_i)(A^{-1}f(x_i))^T \xi(x_i) \\ &= \sum_{i=1}^n A^{-1}f(x_i)f(x_i)^T (A^{-1})^T \xi(x_i) \\ &= A^{-1} \sum_{i=1}^n f(x_i)f(x_i)^T \xi(x_i) (A^{-1})^T \\ &= A^{-1} M_f(\xi) (A^{-1})^T. \end{aligned}$$

Teda

$$\begin{aligned} M_g(\xi) &= A^{-1} M_f(\xi) (A^{-1})^T \\ M_g^{-1}(\xi) &= (A^{-1} M_f(\xi) (A^{-1})^T)^{-1} \\ M_g^{-1}(\xi) &= A^T M_f^{-1}(\xi) A. \end{aligned}$$

□

Pre parametrizáciu modelu (3.4) potrebujeme poznať maticu A , keďže má plnú hodnotu a platí $L = AA^T$, možno ju ľahko nájst napríklad pomocou Choleského rozkladu matice L . Maticu L vypočítame ako $L = \int_{\mathcal{R}} f(x)f(x)^T \mu(dx)$. Výsledný návrh je invariantný na spôsob výpočtu matice A , pretože z vyjadrenia V-optimality ako L-optimality

priamo vyplýva, že pre voľbu A je dôležitá podmienka $L = AA^T$.

Úloha, ktorú riešime má tvar

$$\min_{w \in \mathbb{S}_n} \text{tr}\{A^T M^{-1}(\xi)A\} \Leftrightarrow \min_{w \in \mathbb{S}_n} \text{tr}\{M^{-1}(\xi)AA^T\} \Leftrightarrow \min_{w \in \mathbb{S}_n} \text{tr}\{M^{-1}(\xi)L\}.$$

3.4 Eficiencia A-optimálneho a V-optimálneho návrhu

S cieľom merať kvalitu konkrétnych A-optimálnych resp. V-optimálnych návrhov zavedieme definície ich eficiencií (viac v [11]). Toto však vyžaduje splnenie podmienky pozitívnej homogenity z definície 1.8 pre konkávnou verziu A-optimálneho kritéria $\Psi_A(M(\xi)) = m(\text{tr}(M(\xi)^{-1}))^{-1}$ a V-optimálneho kritéria $\Psi_V(M(\xi)) = \text{tr}\{A^T M^{-1}(\xi)A\}$.

Definícia 3.1. *Relatívnu A-eficienciu návrhu ξ vzhľadom na návrh ζ definujeme ako*

$$\text{eff}_A(\xi|\zeta) = \frac{\text{tr}(M^{-1}(\zeta))}{\text{tr}(M^{-1}(\xi))}.$$

Definícia 3.2. *Relatívnu V-eficienciu návrhu ξ vzhľadom na návrh ζ definujeme ako*

$$\text{eff}_V(\xi|\zeta) = \frac{\text{tr}(A^T M^{-1}(\zeta)A)}{\text{tr}(A^T M^{-1}(\xi)A)}.$$

Pozitívna homogenita oboch kritérií je zrejماً vďaka vlastnosti stopy matice

$$\text{tr}\{\alpha X\} = \alpha \text{tr}\{X\},$$

kde $\alpha \in R$ a X je štvorcová matica.

3.5 Formulácia V-optimality ako SDP

Ukázali sme si, že V-optimality je za istých podmienok možné formulovať ako L-optimality a následne pomocou parametrizácie modelu formulovať ako A-optimality. Úloha SDP, ktorú budeme riešiť je modifikáciou (2.4) a bude vyzeráť

$$\begin{aligned} \min \quad & \sum_{i=1}^m t_i \\ & \begin{pmatrix} \sum_{i=1}^n w_i A^{-1} f(x_i) f(x_i)^T (A^{-1})^T & e_i \\ e_i^T & t_i \end{pmatrix} \succeq 0, \quad i = 1, \dots, m \\ & w \in \mathbb{S}_n. \end{aligned} \quad (3.6)$$

Takáto ekvivalencia je užitočná, keďže existujú dostupné algoritmy na hľadanie A-optimálneho návrhu napr. pomocou SDP a ako sme už spomínali v podkapitole 2.4, v takomto prípade je možné zakomponovať do úlohy dodatočné konvexné ohraničenia na návrh.

3.6 Parametrizácia response surface modelu

Butler v [5] uvádza, že vďaka jednoduchým parametrizáciám A-optimálne kritérium a V-optimálne kritérium sú pre polynomický response surface model ekvivalentné, čo je špeciálny prípad parametrizácie, ktorú sme ukázali my v podkapitole 3.3 pre $\mathcal{R} = \mathcal{X}$ a $\mu(x)$ rovnomernú mieru.

Majme polynomický response surface model stupňa M s k faktormi

$$y(x) = \sum_{r \in S} \beta_r x_1^{r_1} x_2^{r_2} \dots x_k^{r_k} = \sum_{r \in S} \beta_r f_r(x), \quad (3.7)$$

kde $y(x)$ je priemerná odozva (mean response) v bode $x = (x_1, x_2, \dots, x_k)^T$, $r = (r_1, r_2, \dots, r_k)^T$ a $S = \{r : r_1 + \dots + r_k \leq M\}$.

O ekvivalencii A-optimálneho a V-optimálneho kritéria po parametrizácii modelu hovorí nasledovná veta [5]:

Veta 3.2. *V-optimálna a A-optimálna sú pre response surface model (3.7) ekvivalentné, ak regresné funkcie $f_r(x)$ pre $r \in S$ sú ortonormálne v priestore $L^2(\mathcal{R}, \mu(x))$ t.j. $\int_{\mathcal{R}} f_j(x) f_k(x) \mu(x) dx = \delta_{jk}$, kde δ_{jk} je rovné jednej ak $j = k$ a inak je rovné nule.*

Veta 3.2 zahŕňa mnoho parametrizácií response surface modelu a jednou z nich je aj Legendrova parametrizácia, ktorá je špeciálnym prípadom pre rovnomerné pravdepodobnostné rozdelenie $\mu(x)$ na kocke $\mathcal{R} = [-1, 1]^k$.

Je známe, že integrál normalizovaných Legendrových polynómov stupňa r , $P_r(x)$, na intervale $[-1, 1]$ s pravdepodobnostným rozdelením $\mu(x) = \frac{1}{2}$ je rovný jednej t.j.

$$\int_{[-1,1]} P_r(x) P_r(x) \frac{1}{2} dx = 1.$$

Špeciálny prípad vety 3.2 pre Legendrove polynómy má nasledovné znenie [5]:

Dôsledok 3.1. *Pre rovnomerné pravdepodobnostné rozdelenie $\mu(x)$ na kocke $[-1, 1]^k$ V-optimálna je po Legendrovej parametrizácii modelu ekvivalentná A-optimálna.*

4 Príklady a aplikácie

V tejto kapitole na niekoľkých príkladoch ilustrujeme výpočet V-optimálnych návrhov pomocou semidefinitného programovania ako napríklad kvadratický model na štvorci, polynomický model, kvadratický model za marginálnych ohraničení a ohraničení na cenu experimentu a central composite design. V niekoľkých príkladoch okrem V-optimálneho návrhu uvedieme pre porovnanie aj A-optimálny návrh. Všetky výpočty boli prevedené pomocou programu SeDuMi, ktorý rieši optimalizačné problémy nad symetrickými kužeľmi, teda aj úlohy SDP [14].

4.1 Legendrova parametrizácia pre kvadratický model na intervale $[-1, 1]$

Majme model

$$y(x) = \beta^T f(x) + \varepsilon = \beta_0 + \beta_1 x + \beta_2 x^2 + \varepsilon.$$

Naším cieľom je overiť, či parametrizácia uvedená vo vete 3.1 s rovnomerným pravdepodobnostným rozdelením na intervale $[-1, 1]$ je totožná s Legendrovou parametrizáciou modelu, pozri [5].

Parametrizácia modelu podľa vety 3.1 má tvar

$$g(x) = A^{-1}f(x) = A^{-1} \begin{pmatrix} 1 \\ x \\ x^2 \end{pmatrix},$$

pričom $L = AA^T$ a $L = \int_{[-1,1]} f(x)f(x)^T \frac{1}{2}(dx)$. Konkrétne:

$$L = \begin{pmatrix} 1 & 0 & 0,334 \\ 0 & 0,334 & 0 \\ 0,334 & 0 & 0,2 \end{pmatrix},$$
$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0,578 & 0 \\ 0,334 & 0 & 0,3 \end{pmatrix},$$
$$A^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1,73 & 0 \\ -1,12 & 0 & 3,35 \end{pmatrix}, \text{ a teda}$$

$$g(x) = \begin{pmatrix} 1 \\ 1,73x \\ 3,35x^2 - 1,12 \end{pmatrix}.$$

Uvedené regresné funkcie $g_1(x)$, $g_2(x)$ a $g_3(x)$ sú na intervale $[-1, 1]$ a pri rovnomernom pravdepodobnostnom rozdelení ortonormálne polynómy vzhľadom na priestor $\mathcal{L}^2([-1, 1], \frac{1}{2})$ teda spĺňajú podmienku

$$\int_{-1}^1 g_i(x)g_j(x)\frac{1}{2}dx = \begin{cases} 0 & \text{ak } i \neq j \\ 1 & \text{ak } i = j. \end{cases}$$

Je známe, že prvé 3 normalizované Legendrove polynómy na intervale $[-1, 1]$ pre rovnomerné pravdepodobnostné rozdelenie $\mu(x)$ majú tvar:

$$P_0(x) = 1, \quad P_1(x) = \sqrt{3}x, \quad P_2(x) = \sqrt{\frac{5}{4}}(3x^2 - 1). \quad (4.1)$$

Po vykonanej parametrizácii modelu, $g(x) = A^{-1}f(x)$, regresné funkcie $g_1(x)$, $g_2(x)$ a $g_3(x)$ predstavujú prvé tri normalizované Legendrove polynómy. Parametrizácia modelu je ekvivalentná s Legendrovou parametrizáciou, teda podľa vety 3.1 V-optimálny návrh modelu na intervale $[-1, 1]$ je po parametrizácii ekvivalentný s A-optimálnym návrhom. Úlohu hľadania V-optimálneho návrhu možno riešiť ako úlohu SDP podľa (3.6).

V-optimálny návrh pri rovnomernej pravdepodobnostnej miere na $[-1, 1]$ pre polynomický model druhého stupňa má tvar

$$\xi^* = \begin{Bmatrix} -1 & 0 & 1 \\ 0,2512 & 0,4977 & 0,2512 \end{Bmatrix}.$$

Výsledok je v súlade s tvrdením pre exaktný návrh, ktoré uvádza [9]:

Veta 4.1. *Nech $N = 4p + q$, $p \in \mathbf{N}$, $q \in \{-1, 0, 1\}$, $N \neq 5$. Potom V-optimálny návrh pre N meraní pre kvadratický model na $[-1, 1]$ je daný:*

$$\begin{Bmatrix} -1 & 0 & 1 \\ p & 2p + q & p \end{Bmatrix}.$$

4.2 Legendrova parametrizácia pre kvadratický model na štvorci $[-1, 1]^2$

Majme model

$$y(x) = \beta_0 + \beta_1x_1 + \beta_2x_2 + \beta_{11}x_1^2 + \beta_{22}x_2^2 + \beta_{12}x_1x_2 + \varepsilon, \quad (4.2)$$

kde $x = \begin{pmatrix} x_1 & x_2 \end{pmatrix}^T$, $x_1, x_2 \in [-1, 1]$. Cieľom je overiť, či parametrizácia modelu uvedená v podkapitole 3.3 na oblasti $\mathcal{R} = [-1, 1]^2$ s rovnomerným pravdepodobnostným rozdelením $\mu(x)$ je ekvivalentná s Legendrovou parametrizáciou modelu. Parametrizácia modelu má tvar:

$$g(x) = A^{-1}f(x) = A^{-1} \begin{pmatrix} 1 \\ x_1 \\ x_2 \\ x_1^2 \\ x_2^2 \\ x_1x_2 \end{pmatrix},$$

pričom $L = AA^T$ a $L = \int_{[-1,1]^2} f(x)f(x)^T \frac{1}{4} dx$. Konkrétne

$$L = \begin{pmatrix} 1 & 0 & 0 & 0,334 & 0,334 & 0 \\ 0 & 0,334 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0,334 & 0 & 0 & 0 \\ 0,334 & 0 & 0 & 0,2 & 0,112 & 0 \\ 0,334 & 0 & 0 & 0,112 & 0,2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0,112 \end{pmatrix},$$

$$A = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0,578 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0,578 & 0 & 0 & 0 \\ 0,334 & 0 & 0 & 0,299 & 0 & 0 \\ 0,334 & 0 & 0 & 0 & 0,299 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0,334 \end{pmatrix},$$

$$A^{-1} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1,73 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1,73 & 0 & 0 & 0 \\ -1,118 & 0 & 0 & 3,347 & 0 & 0 \\ -1,118 & 0 & 0 & 0 & 3,347 & 0 \\ 0 & 0 & 0 & 0 & 0 & 2,994 \end{pmatrix}, \text{ a teda}$$

$$g(x) = \begin{pmatrix} 1 \\ 1, 73x_1 \\ 1, 73x_2 \\ 3, 35x_1^2 - 1, 12 \\ 3, 35x_2^2 - 1, 12 \\ 2, 5x_1x_2 \end{pmatrix}.$$

Uvedené polynómy $g_i(x)$ pre $i = 1, \dots, 6$ sú na množine $[-1, 1]^2$ a pri rovnomernom pravdepodobnostnom rozdelení ortonormálne vzhľadom na priestor $\mathcal{L}^2([-1, 1]^2, \frac{1}{4})$, teda

$$\int_{-1}^1 \int_{-1}^1 g_i(x)g_j(x) \frac{1}{4} dx_1 dx_2 = \begin{cases} 0 & \text{ak } i \neq j \\ 1 & \text{ak } i = j. \end{cases}$$

Keďže polynómy sú ortonormálne na $\mathcal{L}^2([-1, 1]^2, \frac{1}{4})$, parametrizácia $g(x) = A^{-1}f(x)$ je v súlade s teóriou parametrizácie response surface modelu na kocke $[-1, 1]^k$ podľa vety 3.2.

Dimenzia tejto úlohy SDP sa dá výrazne redukovať. Presnejšie, ako uvádza Heiligers v [8], majme grupu permutácií a zmien znamienok \mathcal{G} na R^k a grupu $\tilde{\mathcal{G}} = \{Q_g : g \in \mathcal{G}\}$, ktorú tvoria také matice Q_g , že $f(g(x)) = Q_g f(x)$ pre všetky $x \in \mathcal{G}$. Nech množina všetkých návrhov \mathcal{X} je invariantná vzhľadom na \mathcal{G} t.j.

$$g(\mathcal{X}) = \mathcal{X} \quad \text{pre všetky } g \in \mathcal{G}$$

a nech kritérium optimality je invariantné vzhľadom na $\tilde{\mathcal{G}}$ t.j.

$$\Psi(M) = \Psi(QMQ^T) \quad \text{pre všetky } M \in \mathcal{M} \text{ všetky } Q \in \tilde{\mathcal{G}},$$

potom každý prípustný návrh ξ má informačnú maticu invariantnú vzhľadom na $\tilde{\mathcal{G}}$ t.j.

$$QM(\xi)Q^T = M(\xi) \quad \text{všetky } Q \in \tilde{\mathcal{G}}.$$

Každý invariantný návrh kvadratického modelu na k -rozmernej kocke má body suportu definované ako $\bigcup_{i=0}^k E_i$, pričom E_i pozostáva z $\binom{k}{i} 2^i$ vektorov s i súradnicami rovnými ± 1 a $k - i$ súradnicami rovnými 0. V našom prípade to znamená, že optimálny aproximatívny návrh kvadratického response surface modelu na štvorci $[-1, 1]^2$ má body suportu v $\{(x_1, x_2) : x_i \in \{-1, 0, 1\}\}$, a teda pri hľadaní optimálneho návrhu sa stačí obmedziť na tieto body.

Výsledný V-optimálny návrh je zobrazený v tabuľke 1. Pre porovnanie výsledkov uvádzame v tabuľke 2 tiež A-optimálny.

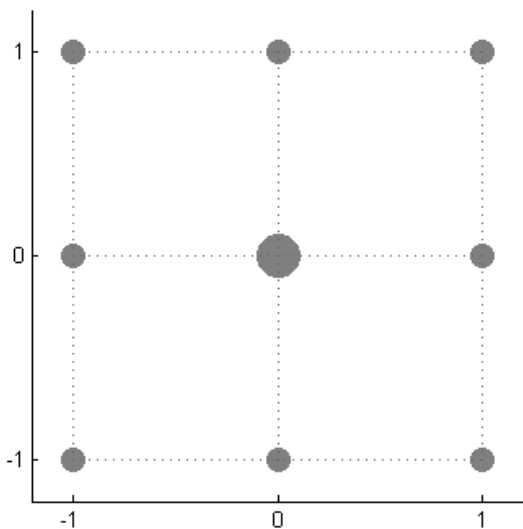
Naše výsledky sú v súlade s výsledným V-optimálnym návrhom pre kvadratický model na štvorci $[-1, 1]^2$, ktorý uvádza [5].

x_i	w_i
$(\pm 1, \pm 1)$	0,0911
$(0, 0)$	0,2703
$(\pm 1, 0)$	0,0913
$(0, \pm 1)$	0,0913

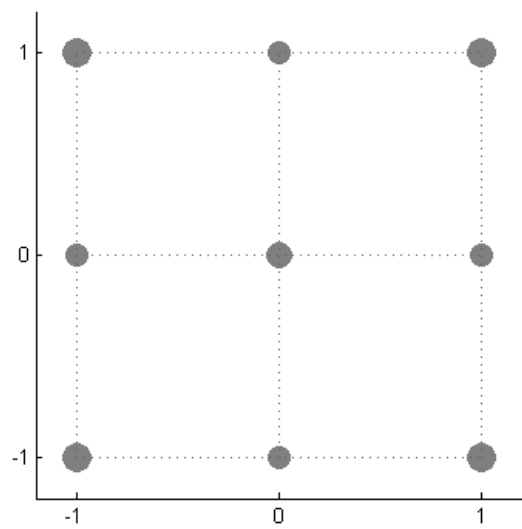
Tabuľka 1: V-optimálny návrh pre kvadratický model na štvorci $[-1, 1]^2$.

x_i	w_i
$(\pm 1, \pm 1)$	0,1288
$(0, 0)$	0,1039
$(\pm 1, 0)$	0,0952
$(0, \pm 1)$	0,0952

Tabuľka 2: A-optimálny návrh pre kvadratický model na štvorci $[-1, 1]^2$.



Obr. 1: V-optimálny návrh pre kvadratický model na štvorci $[-1, 1]^2$



Obr. 2: A-optimálny návrh pre kvadratický model na štvorci $[-1, 1]^2$

4.3 V-optimálny návrh pre polynomický model stupňa d na intervale $[-1, 1]$

Majme polynomický model stupňa d

$$y(x) = \beta_0 + \beta_1 x + \beta_2 x^2 + \dots + \beta_d x^d + \varepsilon. \quad (4.3)$$

Naším cieľom je pre tento model nájsť aproximatívny V-optimálny návrh na oblasti $\mathcal{X} = [-1, 1]$ pomocou SDP tak, ako sme uviedli v podkapitole 3.3. Predpokladajme diskretizáciu intervalu $[-1, 1]$ v n bodoch t.j. $\mathcal{X} = \{-1 = x_1, x_2, \dots, x_n = 1\}$ potom úloha, ktorú budeme riešiť, bude mať podľa vzoru (3.6) tvar

$$\begin{aligned} \min \quad & \sum_{i=1}^{d+1} t_i \\ & \left(\begin{array}{ccc} \sum_{i=1}^n w_i A^{-1} f(x_i) f(x_i)^T (A^{-1})^T & e_i & \\ & e_i^T & t_i \end{array} \right) \succeq 0, \quad i = 1, \dots, d+1 \\ & w \in \mathbb{S}_n, \end{aligned}$$

kde $f(x) = (1, x, x^2, \dots, x^d)^T$ je vektor regresných funkcií. Maticu A získame pomocou Choleského rozkladu matice L , ktorá má tvar

$$L = \int_{[-1,1]} f(x) f(x)^T \mu(dx),$$

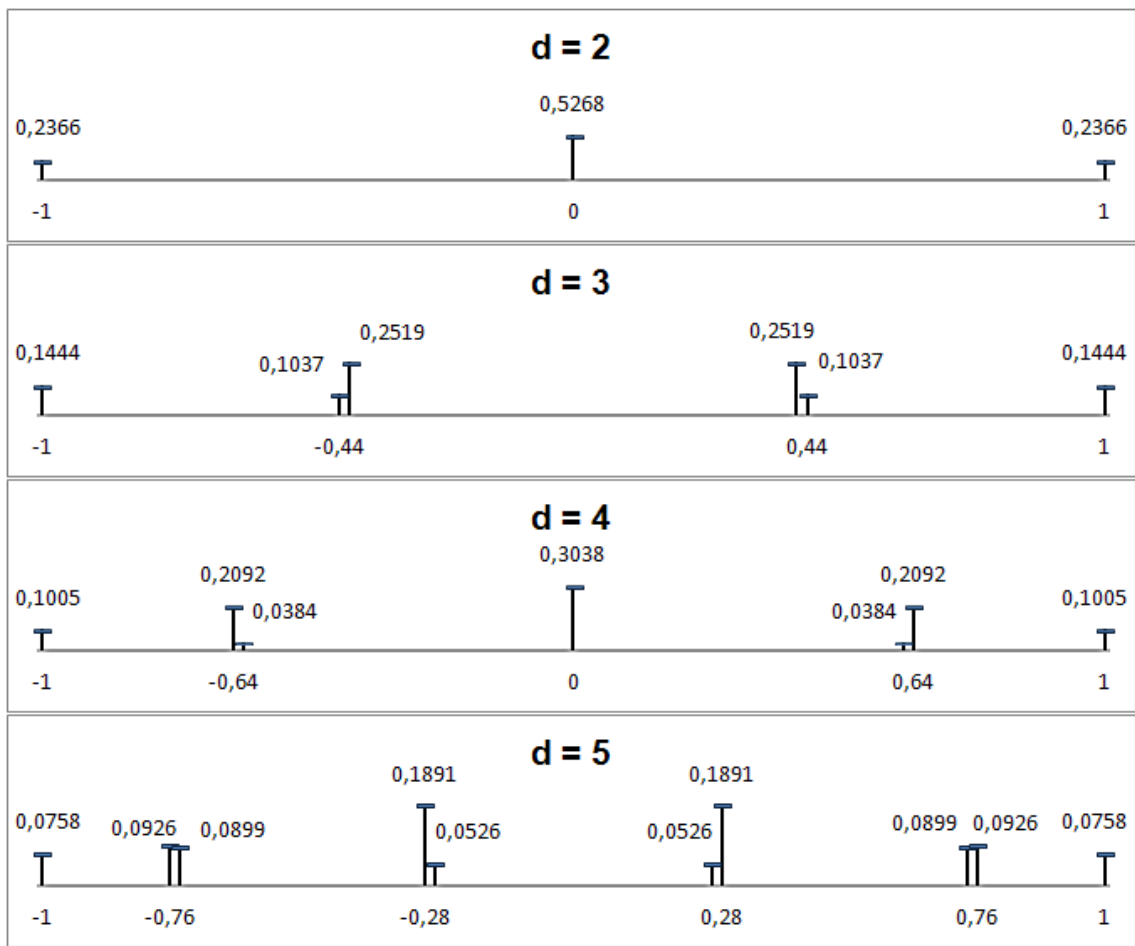
kde $\mu(x)$ je normálne pravdepodobnostné rozdelenie.

Na obrázku 4.3 sú zobrazené V-optimálne návrhy pre polynomický model stupňa 2, 3, 4 a 5 na intervale $[-1, 1]$.

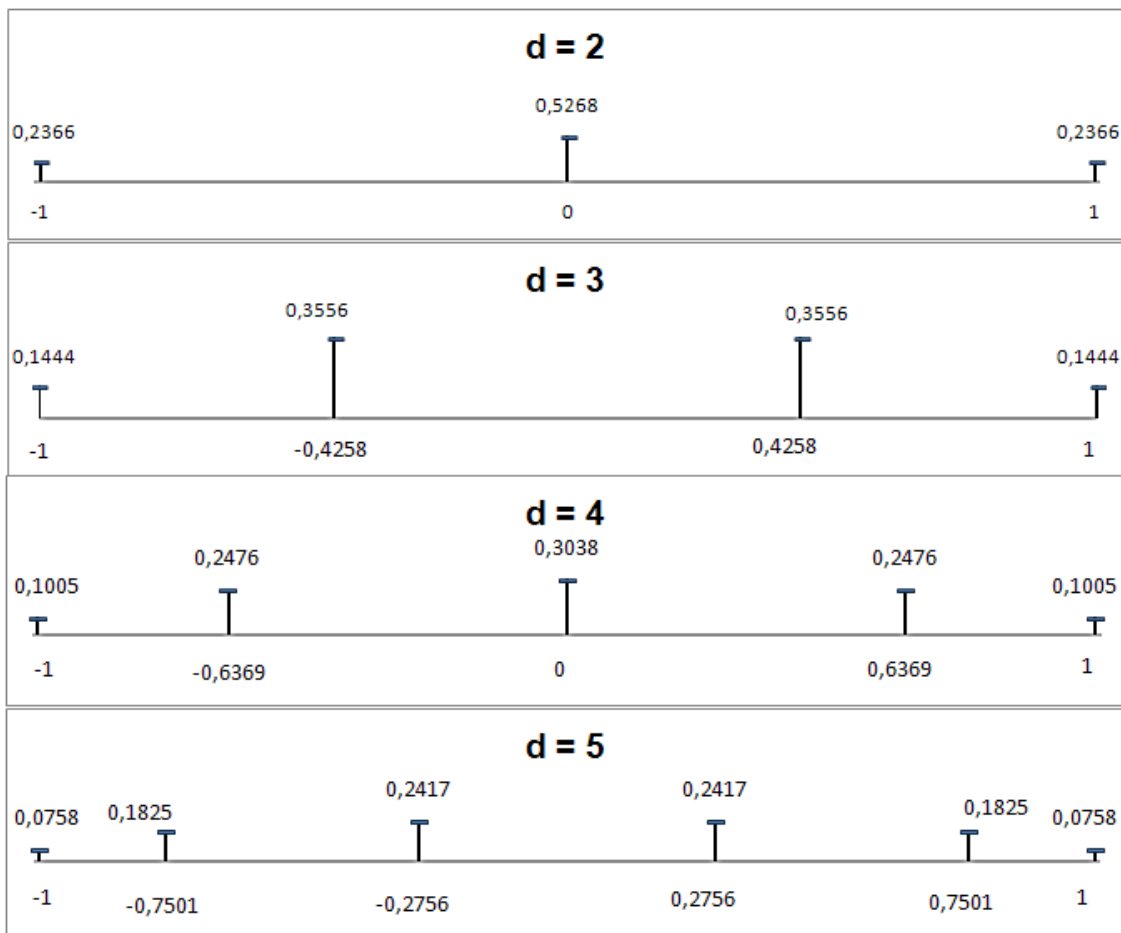
Vďaka tomu, že sme pri výpočte uvažovali diskretizáciu intervalu $[-1, 1]$, vo výslednom návrhu možno pozorovať, že niekoľko bodov suportu je umiestnených tesne vedľa seba (obrázok 4.3, $d = 3, 4, 5$). S cieľom lokalizovať body suportu pre spojitý interval $[-1, 1]$, a nie jeho diskrétnu verziu, nájdeme váženú lineárnu kombináciu týchto párov. Majme dvojice po sebe nasledujúcich bodov a k nim prislúchajúce váhy $\xi(x_i) = w_i$ a $\xi(x_{i+1}) = w_{i+1}$. Pozíciu bodu \bar{x}_i nájdeme ako váženú lineárnu kombináciu

$$\bar{x}_i = \frac{1}{w_i + w_{i+1}} (w_i x_i + w_{i+1} x_{i+1}),$$

pričom nová váha $\xi(\bar{x}_i) = w_i + w_{i+1}$. V-optimálne návrhy pre diskretizáciu $n \rightarrow \infty$ sú zobrazené na obrázku 4.3.



Obr. 3: V-optimální návrh pre polynomický model stupňa d na intervale $[-1, 1]$.



Obr. 4: V-optimálny návrh pre polynomický model stupňa d na intervale $[-1, 1]$, $n \rightarrow \infty$.

4.4 Proces zhusťovania uránového paliva

Majme uránové tyče pripravené na použitie v atómovej elektrárni ako palivo. Každá tyč obsahuje uránové pelety s nejakou špecifickou hustotou. Dôležitou súčasťou procesu je spekanie uránových peliet, ktoré musia mať určitú hustotu, ale aj pórovitosť. Za týmto účelom sa na začiatku procesu pridávajú aditíva, vďaka čomu sa formujú póry na peletách.

Model, s ktorým budeme pracovať je kvadratický s dvoma vysvetľujúcimi premennými

$$y = \beta_0 + \beta_1^2 x_1^2 + \beta_2^2 x_2^2 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2 + \varepsilon.$$

Vektor regresných funkcií bude vyzeráť $f(x) = (1, x_1^2, x_2^2, x_1, x_2, x_1 x_2)^T$.

Pri experimente uvažujeme dve premenné x_1 a x_2 :

1. x_1 je počiatočná hustota v 392 pozorovaných tyčiach od 94,9% do 96,7% , konkrétne $\mathcal{X}_1 = \{94,9; 95,1; 95,2; 95,3; 95,4; 95,5; 95,6; 95,7; 95,8; 95,9; 96,0; 96,1; 96,2; 96,3; 96,4; 96,5; 96,6; 96,7\}$.
2. x_2 predstavuje percento aditíva U_3O_8 , množina $\mathcal{X}_2 = \{0, 10, 20\}$.

Výberový priestor experimentu \mathcal{X} je dvojrozmerná diskretná sieť s rozmermi 18×3 t.j.

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathcal{X} = \begin{pmatrix} \mathcal{X}_1 \\ \mathcal{X}_2 \end{pmatrix}.$$

Proces zhusťovania uránového paliva a jeho aproximatívny D-optimálny návrh hlbšie popisuje [10].

Dodatočné marginálne ohraničenia na počty tyčí

Charakter experimentu si vyžaduje pridanie dodatočných ohraničení na počty tyčí do úlohy. Pre každú tyč s počiatočnou hustotou $d_i, i = 1, \dots, 18$, máme k dispozícii konkrétny počet kusov tyčí n_i . Všetky počty sú uvedené na obrázku 5. Pre nás to znamená zahrnutie dodatočných marginálnych ohraničení na počty kusov tyčí do úlohy. Podľa [3] zahrnutie takýchto dodatočných ohraničení do úlohy SDP je možné zrealizovať veľmi jednoducho. Označme vektor počtov kusov tyčí s počiatočnou hustotou $d_i, i = 1, \dots, 18$ ako $n =$

x_1	94.9	95.1	95.2	95.3	95.4	95.5	95.6	95.7	95.8
n_1	1	3	14	59	52	29	25	32	36
x_1	95.9	96.0	96.1	96.2	96.3	96.4	96.5	96.6	96.7
n_1	29	36	38	12	10	8	2	3	3

Obr. 5: Zastúpenie počiatočných hustôt pre 392 tyčí [10]

$(n_1, n_2, \dots, n_{18})^T$, pričom celkový počet tyčí $N = \sum_{i=1}^{18} n_i = 392$. Máme k dispozícii 54 mrežových bodov, v ktorých možno merať a na výber 18 rôznych počiatočných hustôt tyčí. Každý z hustôt d_i zodpovedajú tri váhy $\xi(d_i, 0)$, $\xi(d_i, 10)$ a $\xi(d_i, 20)$, ktoré určujú akú časť z tyčí s počiatočnou hustotou d_i obohatiť 0%, 10% resp. 20% aditíva. Tieto váhy dokopy musia spĺňať podmienku, že k dispozícii máme presne n_i tyčí s počiatočnou hustotou d_i . Tieto marginálne ohraničenia budú mať tvar

$$\xi(d_i, 0) + \xi(d_i, 10) + \xi(d_i, 20) = \frac{n_i}{N}, \quad i = 1, \dots, 18.$$

Ohraničenia na celkovú cenu experimentu

Vďaka jednoduchosti zahrnúť dodatočné ohraničenia do úlohy SDP sa môžeme zaoberať aj ohraničením na celkovú cenu experimentu. Predpokladajme, že máme obmedzený rozpočet na realizáciu experimentu cenou TC a obohatenie o jedno percento aditíva U_3O_8 stojí jednu peňažnú jednotku. Ohraničenie na celkovú cenu experimentu bude potom vyzeráť

$$10 \sum_{i=1}^{18} N \xi(d_i, 10) + 20 \sum_{i=1}^{18} N \xi(d_i, 20) \leq TC.$$

Hľadanie A-optimálneho návrhu aj s dodatočným ohraničením na celkovú cenu experimentu vedie na úlohu SDP

$$\begin{aligned} \min \quad & \sum_{i=1}^6 t_i \\ & \begin{pmatrix} M(\xi) & e_i \\ e_i^T & t_i \end{pmatrix} \succeq 0, \quad i = 1, \dots, 6 \\ & \xi(d_i, x_2) \geq 0, \quad i = 1, \dots, 54, \quad x_2 \in \{0, 10, 20\} \\ & \sum_{x_2 \in \{0, 10, 20\}} \sum_{i=1}^{18} \xi(d_i, x_2) = 1 \\ & \xi(d_i, 0) + \xi(d_i, 10) + \xi(d_i, 20) = \frac{n_i}{N}, \quad i = 1, \dots, 18 \\ & 10 \sum_{i=1}^{18} N \xi(d_i, 10) + 20 \sum_{i=1}^{18} N \xi(d_i, 20) \leq TC \end{aligned} \tag{4.4}$$

Kvôli numerickým problémom bolo nutné urobiť nasledovnú transformáciu dát

$$\begin{aligned} x_1 : 0 &\rightarrow -1 \\ &10 \rightarrow 0 \\ &20 \rightarrow 1 \end{aligned}$$

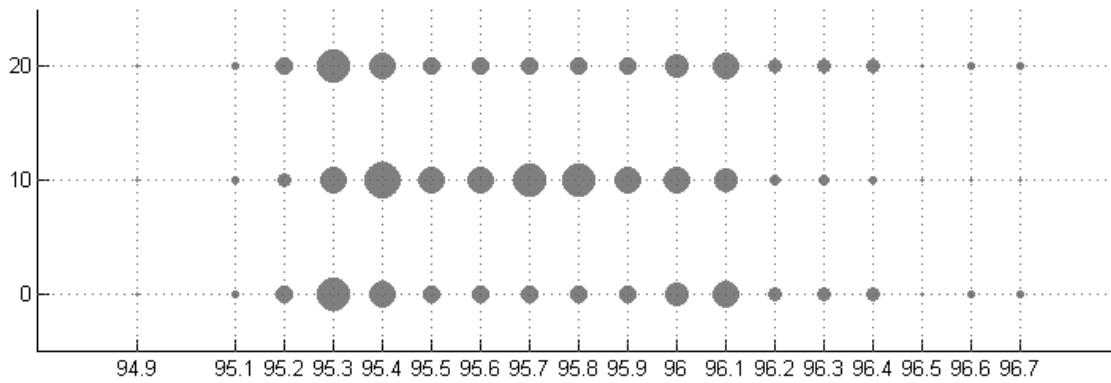
$$x_2(i) = \frac{x_2(i) - \hat{E}(x_2)}{\sqrt{\widehat{Var}(x_2)}}.$$

Na obrázkoch 6, 7 a 8 máme znázornené aproximatívne A-optimálne návrhy pre úlohu (4.4) na dvojrozmernej sieti pre rôzne TC . Veľkosť sivých diskov je úmerná veľkosti váh v daných uzlových bodoch. Návrh, ktorého merania sú približne rovnomerne rozdelené medzi obohatenie o 20%, 10% a 0% aditíva pre počiatočnú hustotu d_i , postupne so znižujúcou sa hranicou sa celkovú cenu experimentu zvyšuje váhy v „lacnejších“ uzloch na úkor uzlov, kde sa obohacuje o 20%. Pri znížení celkovej ceny experimentu o 70% z pôvodnej ceny 3930 sú najväčšie váhy v uzloch, kde sa nebohacuje, čo je najlepšie pozorovateľné hlavne v uzloch, kde máme k dispozícii väčší počet tyčí ako napríklad v uzloch s $d_4 = 95, 3$ a $d_5 = 95, 4$. V uzloch, kde nie je k dispozícii taký veľký počet tyčí $d_2 = 95, 1$ sú váhy rozvrhnuté rovnomerne.

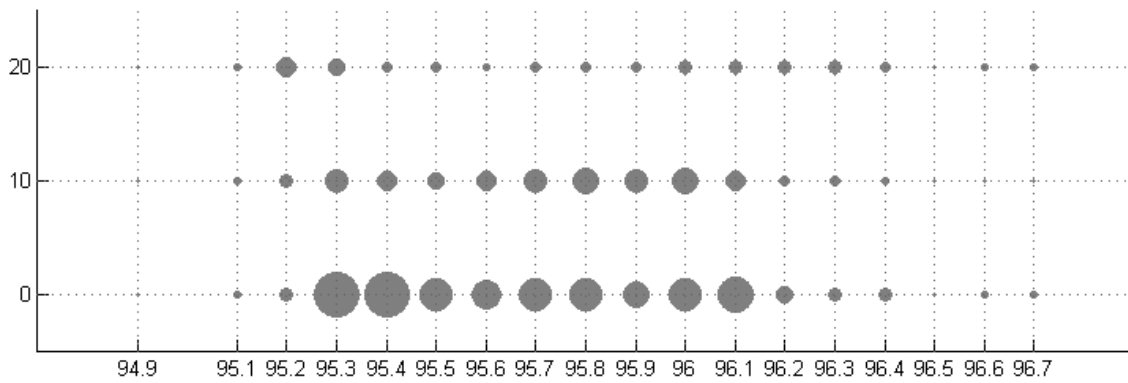
Návrh bez ohraničení na celkovú cenu experimentu je zhodný s návrhom s ohraničením na celkovú cenu experimentu $TC = 3930$. Takáto cena experimentu je najnižšia možná, s akou dosiahneme maximálnu informáciu. V prípade zníženia tejto hranice o polovicu, $TC = 1965$, je efektivita návrhu vzhľadom na A-optimálny návrh bez ohraničenia na cenu 0,78, teda polovičné náklady na experimentu nás stoja stratu 22% informácie. Pri znížení ceny experimentu o 70% A-optimálny návrh dosahuje efektivitu vzhľadom na A-optimálny návrh bez ohraničenia na cenu 0,57. Pri redukcii nákladov na experiment o 70% sa stále zachováva viac ako 50% informácie.

Hľadanie V-optimálneho návrhu vďaka parametrizácii úlohy (4.4) vedie opäť na úlohu SDP

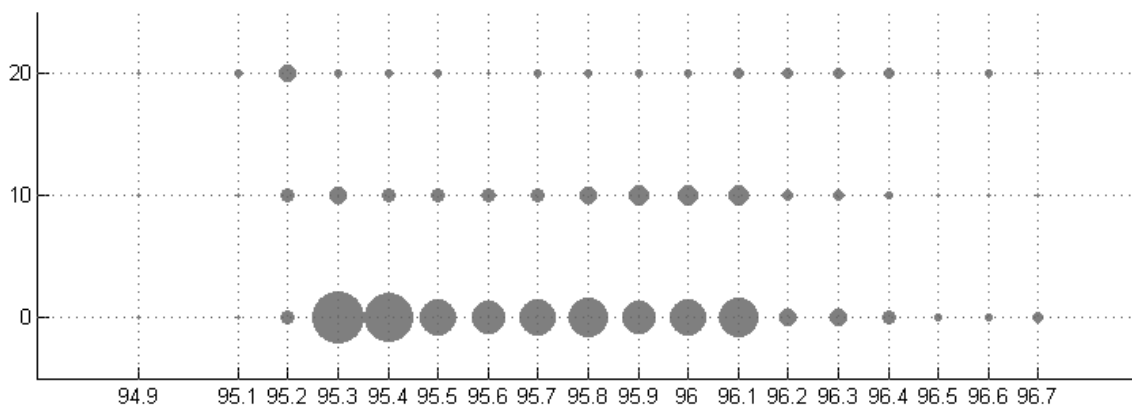
$$\begin{aligned}
\min \quad & \sum_{i=1}^6 t_i \\
& \begin{pmatrix} A^{-1}M(\xi)(A^{-1})^T & e_i \\ e_i^T & t_i \end{pmatrix} \succeq 0, \quad i = 1, \dots, 6 \\
& \xi(d_i, x_2) \geq 0, \quad i = 1, \dots, 54, \quad x_2 \in \{0, 10, 20\} \\
& \sum_{x_2 \in \{0, 10, 20\}} \sum_{i=1}^{18} \xi(d_i, x_2) = 1 \\
& \xi(d_i, 0) + \xi(d_i, 10) + \xi(d_i, 20) = \frac{n_i}{N}, \quad i = 1, \dots, 18 \\
& 10 \sum_{i=1}^{18} N \xi(d_i, 10) + 20 \sum_{i=1}^{18} N \xi(d_i, 20) \leq TC
\end{aligned} \tag{4.5}$$



Obr. 6: Výsledný A-optimálny návrh pre úlohu (4.4) s ohraničením na celkovú cenu experimentu $TC = 3930$



Obr. 7: Výsledný A-optimálny návrh pre úlohu (4.4) s ohraničením na celkovú cenu experimentu $TC = 1965$, redukcia nákladov o 50%



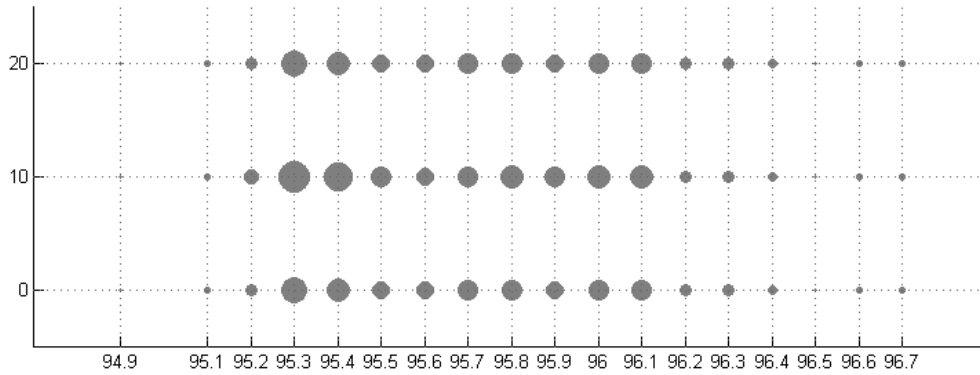
Obr. 8: Výsledný A-optimálny návrh pre úlohu (4.4) s ohraničením na celkovú cenu experimentu $TC = 1179$, redukcia nákladov o 70%

Maticu A získame pomocou Choleského rozkladu matice L , ktorá má tvar

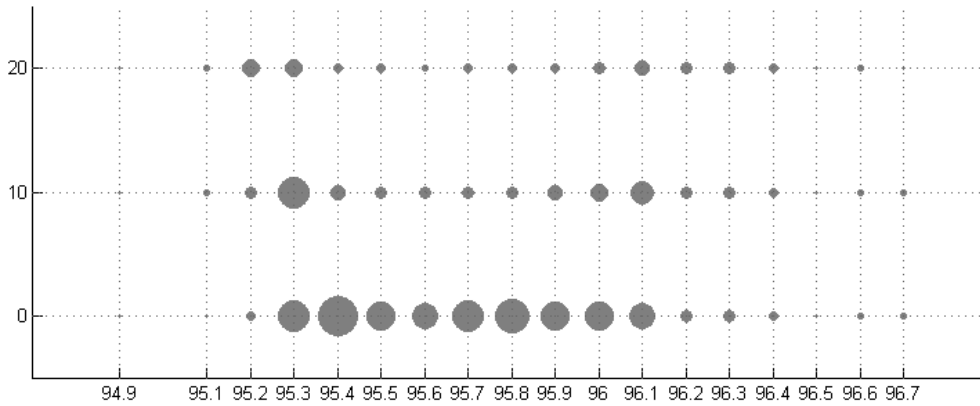
$$L = \int_{\text{conv}(\mathcal{X})} f(x)f(x)^T \mu(dx),$$

kde $\mu(x)$ je rovnomerné pravdepodobnostné rozdelenie. Parametrizácia, ktorú sme vykonali má tvar $g(x) = A^{-1}f(x)$, teda úloha hľadania V-optimálneho návrhu pre pôvodný model s vektorom regresných funkcií $f(x)$ sa po tejto parametrizácii mení na hľadanie A-optimálneho návrhu pre parametrizovaný model s vektorom regresných funkcií $g(x)$, ako uvádzame v (4.5).

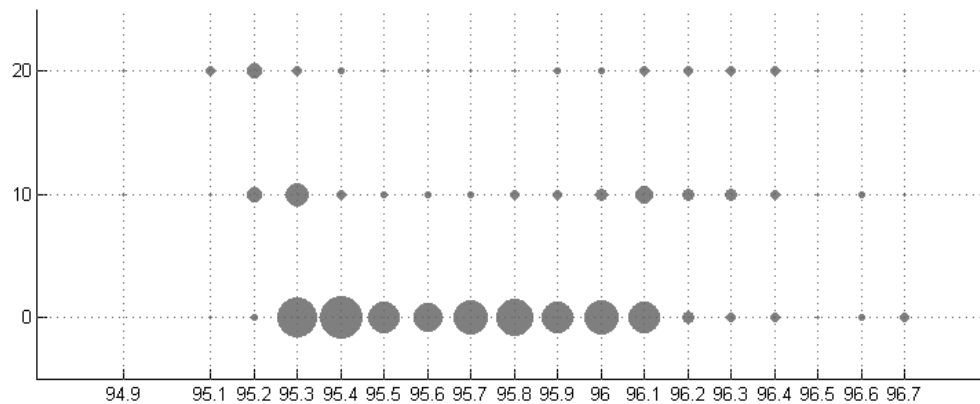
Obrázky 9, 10 a 11 sú grafickým znázornením aproximatívneho V-optimálneho návrhu s ohraničením na celkovú cenu experimentu na 3930, 1965 a 1179 peňažných jednotiek. Podobne ako v prípade A-optimálneho návrhu s ohraničením na celkovú cenu experimentu 3930 aj V-optimálny návrh s takýmto ohraničením je tento návrh zhodný s V-optimálnym návrhom bez ohraničenia na celkovú cenu experimentu. Opäť možno pozorovať jav, že s postupným znižovaním TC sa váhy presúvajú do uzlov, kde obohacovanie aditívom U_3O_8 stojí menej resp. nič. V prípade zníženia ohraničenia na celkovú cenu experimentu o 50% efíciencia vzhľadom na V-optimálny návrh bez ohraničenia na cenu experimentu je rovná 0,86, teda pri úspore nákladov o polovicu sa stále zachováva 86% informácie. Pri znížení ohraničenia na celkovú cenu experimentu na 30% je efíciencia V-optimálneho návrhu vzhľadom na V-optimálny návrh bez ohraničenia na cenu experimentu rovná 0,69.



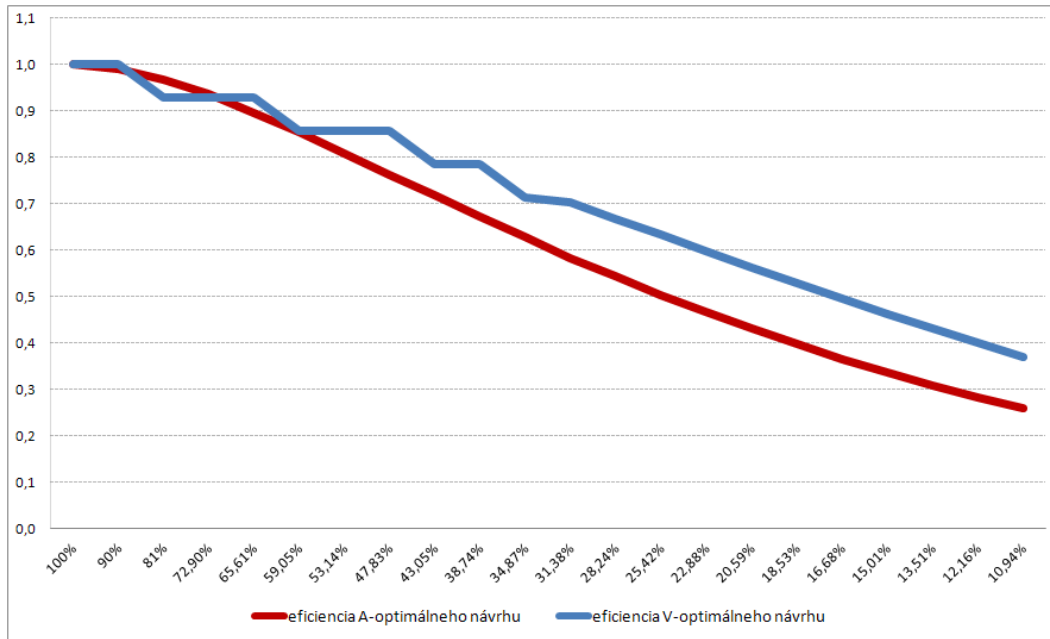
Obr. 9: Výsledný V-optimálny návrh pre úlohu (4.5) s ohraničením na celkovú cenu experimentu $TC = 3930$



Obr. 10: Výsledný V-optimálny návrh pre úlohu (4.5) s ohraničením na celkovú cenu experimentu $TC = 1965$, redukcia nákladov o 50%



Obr. 11: Výsledný V-optimálny návrh pre úlohu (4.5) s ohraničením na celkovú cenu experimentu $TC = 1179$, redukcia nákladov o 70%



Obr. 12: Porovnanie relatívnych A-eficiencií a V-eficiencií (vertikálna os) aproximatívnych návrhov pre rôzne redukcie celkovej ceny experimentu TC (horizontálna os) vzhľadom na V- resp. A-optimálny návrh bez ohraničenia na cenu experimentu.

4.5 Central Composite Design

Central composite design je jeden z najznámejších návrhov druhého stupňa. Jedná sa o návrh, ktorý pozostáva z n_f faktorových bodov (factorial points), z n_0 centrálnych bodov (central points) a n_α osových bodov (axial points). Faktorové body nadobúdajú hodnoty ± 1 pre každý faktor. Osové body sú body umiestnené v určitej vzdialenosti α od stredu návrhu v oboch smeroch každej definovanej osi t.j. ak pracujeme s p faktormi, tak máme k dispozícii $2p$ rôznych osových bodov. Oblasť \mathcal{X} pozostáva z 9-tich bodov $\{(0, 0), (\pm 1, \pm 1), (\pm \alpha, 0), (0, \pm \alpha)\}$.

Uvažujme model s dvoma faktorovými premennými

$$y(x) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{11} x_1^2 + \beta_{22} x_2^2 + \beta_{12} x_1 x_2 + \varepsilon.$$

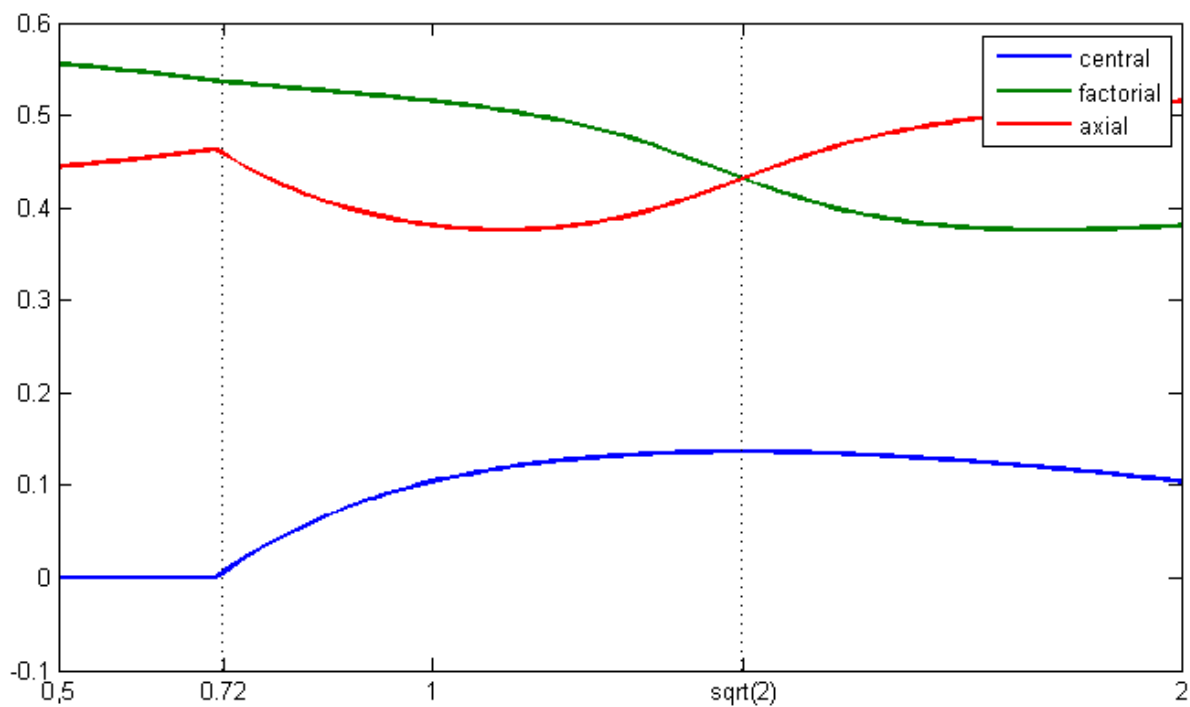
Naším cieľom je zistiť priebeh váh A-optimálneho resp. V-optimálneho návrhu (factorial, central a axial) pre meniacu sa oblasť \mathcal{X} t.j. pre rôzne hodnoty α . Úloha hľadania V-optimálneho návrhu bola riešená po parametrizácii modelu s vektorom regresných funkcií $f(x)$ na $g(x) = A^{-1}f(x)$ pomocou semidefinitného programovania podobne ako v podkapitole 4.4 pre každé α zvlášť. Maticu A potrebnú pre parametrizáciu modelu získame pomocou Choleského rozkladu $L = AA^T$ a $L = \int_{\mathcal{R}} f(x)f(x)^T \mu(dx)$, kde oblasť $\mathcal{R} =$

$[-\max(1, \alpha), \max(1, \alpha)] \times [-\max(1, \alpha), \max(1, \alpha)]$ a $\mu(x)$ berieme rovnomernú a normálnu pravdepodobnostnú mieru na oblasti \mathcal{R} .

Na obrázku 13 si možno všimnúť rovnosť váh zodpovedajúcim axial a factorial v prípade, keď $\alpha = \sqrt{2}$. V tomto prípade návrh reprezentuje 8 bodov rovnomerne rozvrhnutých na kružnici a jeden bod v strede kružnice. Ako uvádza [6], takáto voľba $\alpha = (n_f)^{\frac{1}{4}}$, pričom v našom prípade $n_f = 4$, zabezpečuje rotovateľnosť návrhu (rotatability). Takýto návrh, nesie rovnaké množstvo informácie o odozve y v rovnakej vzdialenosti d v každom smere od stredu návrhu. Výsledný A-optimálny návrh pre $\alpha = \sqrt{2}$ je znázornený v tabuľke 3.

Na obrázku 14 si možno všimnúť, že oproti A-optimálnemu návrhu sa krivky zodpovedajúce priebehu váh vo factorial a axial points pretnú o niečo skôr, a to v $\alpha = 1$. Príslušný V-optimálny návrh pre $\alpha = 1$ možno nájsť v tabuľke 4.

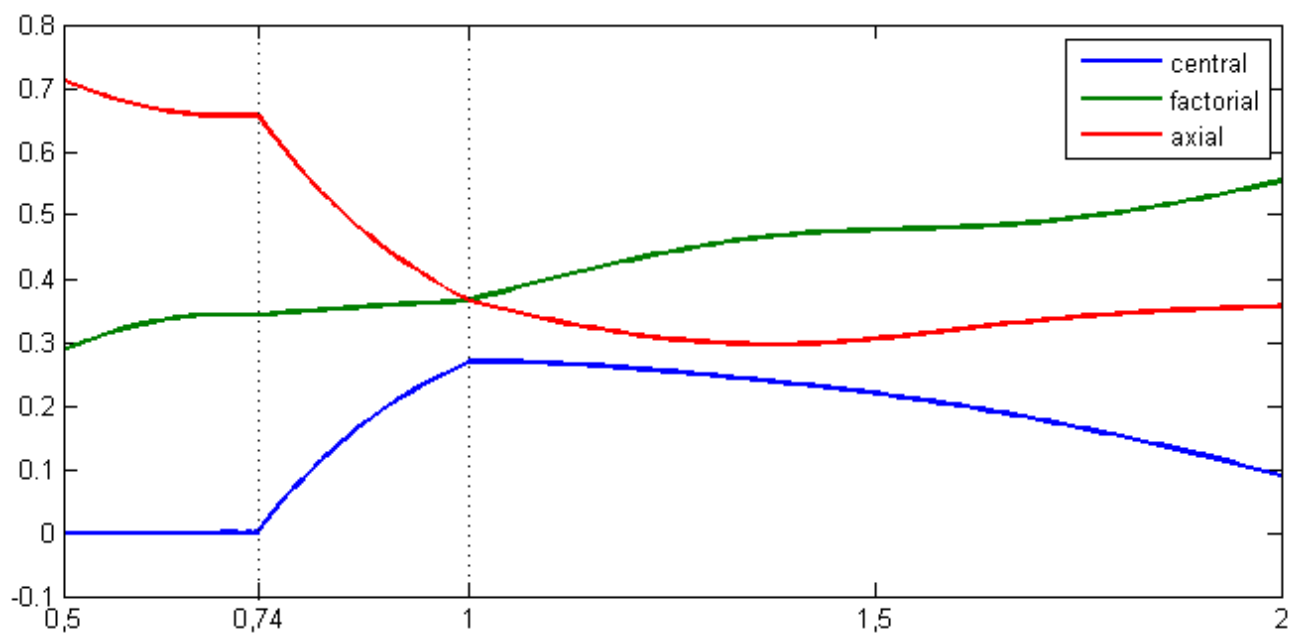
Na obrázku 15 je zobrazený priebeh váh V-optimálneho návrhu vzhľadom na meniacu sa oblasť \mathcal{R} pre normálnu pravdepodobnostnú mieru. Možno si všimnúť, že určité charakteristiky priebehu sa zachovali ako napríklad rovnosť váh vo factorial a axial points je tiež približne 1 points a podobne vyzerá aj zmena priebehu stredového bodu na nenulovú hodnotu. Návrh pre $\alpha = 1$ možno nájsť v tabuľke 5.



Obr. 13: Priebch váh A-optimálneho návrhu pre central composite design s dvoma faktorovými premennými vzhľadom na meniace sa α .

$\alpha = \sqrt{2}$	
central	0,1365
factorial	0,1079
axial	0,1079

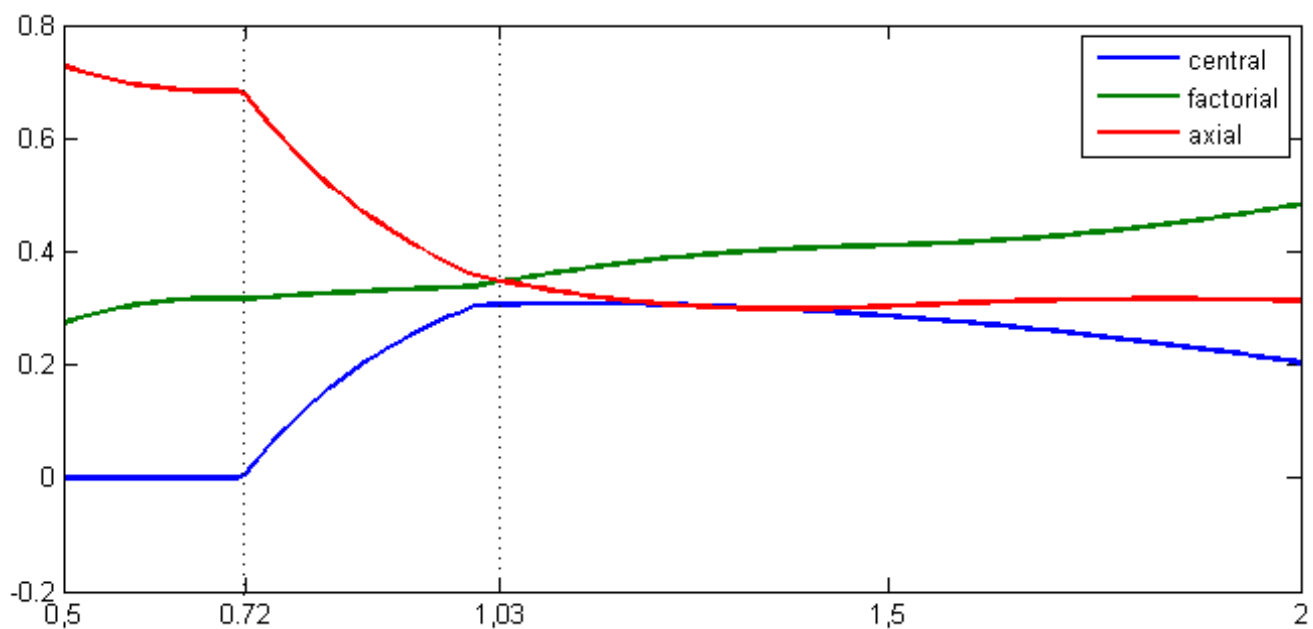
Tabuľka 3: A-optimálny návrh pre $\alpha = \sqrt{2}$.



Obr. 14: Priebeh váh V-optimálneho návrhu pre rovnomernú mieru pre central composite design s dvoma faktorovými premennými vzhľadom na meniace sa α .

$\alpha = 1$	
central	0,2697
factorial	0,0912
axial	0,0914

Tabuľka 4: V-optimálny návrh pre $\alpha = 1$.



Obr. 15: Priebeh váh V-optimálneho návrhu pre normálnu mieru pre central composite design s dvoma faktorovými premennými vzhľadom na meniace sa α .

$\alpha = 1,03$	
central	0,306
factorial	0,0866
axial	0,087

Tabuľka 5: V-optimálny návrh pre $\alpha = 1,03$.

Záver

Cieľom práce bolo aplikovať metódy semidefinitného programovania v optimálnom navrhovaní experimentov, a to najmä v tých situáciách, kde klasické prístupy výpočtov optimálnych návrhov zlyhávajú. Tieto postupy sme ilustrovali na niekoľkých lineárnych modeloch, kde úlohou bol výpočet V-optimálneho alebo A-optimálneho návrhu s možnými dodatočnými lineárnymi ohraničeniami na návrh. Všetky výpočty boli prevedené pomocou softvéru SeDuMi, ktorý rieši optimalizačné problémy nad symetrickými kužeľmi, teda aj problémy SDP.

Zamerali sme sa predovšetkým na hľadanie V-optimálnych návrhov pomocou SDP. Ukázali sme, že po vhodnej parametrizácii pôvodného modelu možno kritérium V-optimality formulovať ako kritérium A-optimality v novom modeli a riešiť ako úlohu SDP. Parametrizácia vyplývala z formulácie V-optimality ako L-optimality.

Overili sme, že hľadanie V-optimálneho návrhu pre kvadratický model na $[-1, 1]$ po vykonaní Legendrovej parametrizácie modelu je ekvivalentné s hľadaním A-optimálneho návrhu. Výsledný V-optimálny návrh pre kvadratický model na štvorci $[-1, 1]^2$ sme hľadali ako A-optimálny návrh po vykonaní Legendrovej parametrizácie na oblasti $[-1, 1]^2$. Výsledky V-optimálneho návrhu pre kvadratický model na štvorci sú známe a sú v súlade s našimi. Pri hľadaní V-optimálnych návrhov pre polynomický model stupňa $d = 2, 3, 4, 5$ na intervale $[-1, 1]$ v dôsledku diskretizácie intervalu v n bolo možné pozorovať, že niekoľko bodov suportu bolo umiestnených tesne vedľa seba. S cieľom lokalizovať body suportu pre $n \rightarrow \infty$ sme použili váženú lineárnu kombináciu týchto bodov.

Prihliadnuc k interpretácii kritéria V-optimality, najdôležitejšou aplikáciou bolo hľadanie V-optimálneho návrhu pre kvadratický response surface model. Charakter experimentu si vyžadoval do úlohy SDP zahrnúť dodatočné marginálne ohraničenia a ohraničenie na celkovú cenu experimentu. Výpočty sme realizovali pre rôzne ohraničenia na celkovú cenu experimentu a výsledky porovnali s príslušnými A-optimálnymi návrhmi. Takisto sme porovnali relatívne efencie V-optimálnych a A-optimálnych návrhov pre rôzne redukcie celkovej ceny experimentu vzhľadom na príslušné návrhy bez ohraničení na celkovú cenu experimentu. Poslednou aplikáciou bol výpočet V-optimálneho návrhu pre central composite design pre dva faktory. Tu nás zaujímal najmä priebeh váh vzhľadom na postupne meniacu sa oblasť \mathcal{X} pre V-optimálny, ale aj A-optimálny návrh, kde v priebehoch možno pozorovať odlišnosti.

Literatúra

- [1] Alizadeh F.: *Interior point methods in semidefinite programming with applications to combinatorial optimization*, SIAM Journal on Optimization 5 (1993), 13-51.
- [2] Atkinson A.C., Donev A.N., Tobias R.D.: *Optimum experimental designs, with SAS*, Oxford University Press, New York (2007).
- [3] Boyd S., Vandenberghe L.: *Application of semidefinite programming*, Applied Numerical Mathematics 29 (1999), 283-299.
- [4] Boyd S., Vandenberghe L.: *Semidefinite programming*, SIAM Review 38 (1996), 49-95.
- [5] Butler N.: *On weighted design optimality criteria and response surface parametrizations*, Statistics & Probability Letters 51 (2001), 41-46.
- [6] Dean A., Voss D.: *Design and Analysis of Experiments*, Springer, New York (1999).
- [7] Harman R., Jurík T.: *Computing c-optimal experimental designs using the simplex method of linear programming*, Computational Statistics & Data Analysis 53 (2008), 247-254.
- [8] Heiligers B.: *Admissible experimental designs in multiple polynomial regression*, Journal of Statistical Planning and Inference 31 (1992), 219-233.
- [9] Imhof L.: *Exact design minimising the integrated variance in quadratic regression*, Statistics: A Journal of Theoretical and Applied Statistics 34 (2000), 103-115.
- [10] Martín-Martín R., Torsney B., López-Fidalgo J.: *Construction of marginally and conditionally restricted designs using multiplicative algorithms*, Computational Statistics & Data Analysis 51 (2007), 5547-5561.
- [11] Pukelsheim F.: *Optimal Design of Experiments*, John Wiley & Sons, New York (1993).
- [12] Pukelsheim F., Rieder S.: *Efficient rounding of approximate designs*, Biometrika 79 (1992), 763-770.
- [13] Saigal R., Vandenberghe L., Wolkowicz H., editors: *Handbook of semidefinite programming: Theory, algorithms and applications*, Kluwer's Academic Publishers (2003).

- [14] Sturm J.F.: *Using SEDUMI 1.02, A MATLAB toolbox for optimization over symmetric cones*, Optimization methods and software 11 (1999), 625-653.

Príloha

Proces zhustovania uránového paliva: V-optimálny návrh s ohraňením na celkovú cenu experimentu

```
function Vopt_uran_tc(TC)
%vstup TC je celkova cena experimentu
%A optimalny navrh pre parametrizovany model g=inv(A)*f

n1 = [1 3 14 59 52 29 25 32 36 29 36 38 12 10 8 2 3 3]; %pocety tyci
n2 = sum(n1); %celkovy pocet tyci
n3 = length(n1);
d = 5;
%siet
x = [94.9 94.9 94.9 95.1 95.1 95.1 95.2 95.2 95.2 95.3 95.3 95.3 95.4 95.4...
     95.4 95.5 95.5 95.5 95.6 95.6 95.6 95.7 95.7 95.7 95.8 95.8 95.8 95.9...
     95.9 95.996 96 96 96.1 96.1 96.1 96.2 96.2 96.2 96.3 96.3 96.3 96.4...
     96.4 96.4 96.5 96.5 96.5 96.6 96.6 96.6 96.7 96.7 96.7; 0 10 20 0 10...
     20 0 10 20 0 10 20 0 10 20 0 10 20 0 10 20 0 10 20 0 10 20 0 10...
     10 20 0 10 20 0 10 20 0 10 20 0 10 20 0 10 20 0 10 20 0 10 20];
x = x';
n = size(x,1);
%ceny
cost = x(:,2);
%normalizacia dat
e = mean(x(:,1));
s = std(x(:,1));
xn = (x(:,1)-e)/s;
xn = [xn, [-1 0 1 -1 0 1 -1 0 1 -1 0 1 -1 0 1 -1 0 1 -1 0 1 -1...
          0 1 -1 0 1 -1 0 1 -1 0 1 -1 0 1 -1 0 1 -1 0 1 -1 0 1]'];

%vypocet matice A, L=AA^t, roznomerna pravdepodobnostna miera,
%R je obdlznik
L = zeros(d+1, d+1);
Li = zeros(d+1, d+1);
r1 = xn(1, 1):0.01:xn(n,1);
r2 = -1:0.1:1;
dlzka_r1 = length(r1);
```

```

dlzka_r2 = length(r2);

for i = 1:dlzka_r1
    for j = 1:dlzka_r2
        Li = (1/dlzka_r1*dlzka_r2)*[1, r1(i)^2, r2(j)^2, r1(i),...
            r2(j), r1(i)*r2(j)]'*[1, r1(i)^2, r2(j)^2, r1(i),...
            r2(j), r1(i)*r2(j)];
        L = L + Li;
    end
end

%cholesky
A = chol(L, 'lower');
A_inv=inv(A);

C = sparse(zeros((d+2)*(d+2), d+1));
Ai = sparse(zeros( ((d+1)*(d+2)+n+2+2*n3+1)^2,d+1+n) );
H = sparse(zeros((d+1)*(d+1), d+1));
E = sparse(zeros(n*n, n));
H_pom = sparse(zeros((d+2)*(d+2),d+1));
POM = sparse(zeros((d+2)*(d+2),1));
D = sparse(zeros((d+1)*(d+2)*(d+1)*(d+2), d+1));
M = sparse(zeros(n3^2, n));

%vektor b
b = sparse(-[ones(d+1,1); zeros(n,1)]);

%matica C
C_w1 = sparse(zeros(n));
C_w2 = sparse([1:2], [1:2], [1, -1]);
for i = 1:d+1
    C(:,i) = vec(sparse([i,d+2],[d+2,i],1));
end
C_diag = mat(C(:,1));
for i = 2:d+1
    C_diag = sparse(blkdiag(C_diag, mat(C(:,i))));
end
C_w3 = sparse(diag(n1)/n2);

```

```

C_diag = sparse(blkdiag(C_diag, C_w1, C_w2, C_w3, -C_w3, TC));
%vytvorenie vektora c
c = vec(C_diag');

%vytvorenie matic Ai
%Ai pre ti
for i = 1:d+1
    Ai(((d+1)*(d+2)+n+2+2*n3+1)*((d+1)*i + i-1) + (d+2)*i , i)=-1;
end
%Ai pre wi
for i = 1:n
    H(:,i) = vec(-A_inv*[1, xn(i,1)^2, xn(i, 2)^2, xn(i,1),...
        xn(i,2), xn(i,1)*xn(i,2)]'*[1, xn(i,1)^2,...
        xn(i, 2)^2, xn(i,1), xn(i,2), xn(i,1)*xn(i,2)]*(A_inv)' );
    E(:,i) = vec(sparse(i,i,1,n,n));
    H_pom(:,i) = vec(sparse([mat(H(:,i)),zeros(d+1,1); zeros(1,d+2)]));
    M(:,i) = vec(sparse(ceil(i/3), ceil(i/3), 1, n3, n3));
end

for i = 1:n
    POM = mat(H_pom(:,i));
    for j = 2:d+1
        POM = sparse(blkdiag(POM, mat(H_pom(:,i))));
    end
    D(:, i) = vec(POM);
end

for i = 1:n
    Ai(:,d+1+i) = vec(sparse(blkdiag(mat(D(:,i)), mat(-E(:,i)), C_w2, ...
        mat(M(:, i)), mat(-M(:,i)), cost(i)*n2)));
end

Ai = Ai';
K.s = size(C_diag',1);
[xt, y, info] = sedumi(Ai, b, c, K);
wi = y(d+2:d+1+n);
ksi = [x, wi]

```