



Силви-Мария Тодорова Гюрова

Стохастични числени методи за оценка на
собствени стойности

АВТОРЕФЕРАТ

на Дисертация

за придобиване на образователна и научна степен “доктор”

по докторска програма: “Математическо моделиране и приложение на математиката”

научна област: 4. “Природни науки, математика и информатика”

професионално направление: 4.5. “Математика”

научен ръководител: проф. д-р Анета Недева Караиванова

Дисертационният труд е обсъден и допуснат до защита на разширено заседание на секция „Скалируеми алгоритми и киберсигурност с Център по високопроизводителни пресмятания“ на ИИКТ-БАН, състояло се на **25.03.2026 г.** от **14:00** часа в зала **2** на блок 25А на ИИКТ-БАН.

Дисертацията съдържа **118** страници и включва **15** фигури и **16** таблици, както и списък с математически означения и използвани абривиатури. Дисертационният труд е структуриран в увод, три основни глави, заключение, списък на публикациите, апробация на резултатите, формулировка на научните и научно-приложните приноси, приложение и **10** страници библиография, включваща **121** заглавия.

Защитата на дисертацията ще се състои наг. от часа в зала **218** на блок 25А на ИИКТ-БАН на открито заседание с научно жури в състав:

Научно жури:

1.
2.
3.
4.
5.

Рецензиите и становищата на членовете на научното жури и авторефератът са публикувани на сайта на ИИКТ-БАН.

Материалите за защитата са на разположение за интересувалите се на адрес: ИИКТ-БАН, ул. Акад. Г. Бончев, блок 2, ет. 3, стая

Автор: **Силви-Мария Тодорова Гюрова**

Заглавие: **Стохастични числени методи за оценка на собствени стойности**

Обща характеристика на дисертационния труд

Актуалност на темата

Задачата за намиране и оценяване на екстремалните собствени стойности на квадратни, в частност симетрични матрици, заема централно място в линейната алгебра поради ключовата си роля в широк кръг научни и приложни области. Такива задачи възникват в квантовата механика, при уравнението на Шрьодингер [65], собствените стойности на Хамилтоновия оператор определят енергийния спектър на системата, като екстремалните стойности дават информация за основното и възбудените състояния [64, 73]. Във финансовата математика анализът на екстремалните собствени стойности на ковариационните матрици е съществен за моделиране на риска и оптимизация на портфейли, включително за оценка на систематичния пазарен риск и възможностите за диверсификация [51, 72]. В спектралната теория на графите тези стойности са свързани със структурните свойства като свързаност, разширение и устойчивост, които са определящи за ефективността на алгоритми в машинното обучение и мрежовия анализ [8, 12, 26]. В обработката на изображения и компютърното зрение собствените стойности играят ключова роля при анализа на главните компоненти, редукцията на размерността, потискането на шума и спектралната сегментация [34, 47]. В тензорния анализ и обработката на многомерни данни, максималната собствена стойност и спектралните норми определят поведението на итеративните методи и техниките за регуляризация, с приложения в обработката на сигнали и изкуствения интелект [49, 56, 80].

Съществуват различни числени подходи за изчисляване на собствени стойности и вектори - точни (директни), итерационни и стохастични. Към директните методи спадат QR декомпозицията и решаването на характеристичния полином при матрици с малка размерност. Те намират решението за краен брой стъпки и имат изчислителна сложност от порядък $\mathcal{O}(n^3)$, където n е размерността на матрицата. Основен недостатък на точните методи е, че при матрици с големи размерности, изчисленията с крайна точност водят до закръгляване и натрупване на грешки, което може да наруши точността на собствените стойности. [87, 90].

Итерационните методи, като например Степенния метод, Обратния Степенен метод, метод на Якоби, метод на Арнолди и метод на Ланцош [33, 43], имат по-малка изчислителна сложност $\mathcal{O}(n^2k)$, където n е размерността на матрицата, а k е броят на итерациите. Този тип методи често се характеризират с бавна сходимост, силна зависимост от началното приближение, значителни изисквания към паметта и ограничена ефективност при паралелизация.

Стохастичните числени методи предлагат ефективна алтернатива, като осигуряват линейна зависимост между размерността на задачата и необходимата памет, както и естествен паралелизъм и относителна лекота на реализация. Макар и с по-бавна сходимост и по-малка прецизност в сравнение с детерминистичните методи, те позволяват решаването на задачи с много големи размерности и са особено подходящи за оценяване на екстремални собствени стойности на симетрични матрици.

В условията на бързото развитие на високопроизводителните изчислителни системи (англ. *High-Performance Computing*, HPC) и навлизането на нови хардуерни архитектури нараства необходимостта от разработване и внедряване на софтуерни решения, базирани на ефективни стохастични алгоритми за задачи с висока размерност, включително за оценяване на екстремални собствени стойности [18, 24, 59]. Това обуславя засилен изследователски интерес към повишаване на точността, ускоряване на сходимостта и разширяване на приложимостта на стохастичните методи в контекста на научните изчисления, машинното обучение и анализа на данни, съобразно съвременните изчислителни възможности.

Обзор на основните резултати в областта

Първите публикации, свързани с Монте Карло (МК) методите в линейната алгебра, датират от 50-те и 60-те години на ХХ век, като първоначално не е разглеждан пряко проблемът за приближено изчисляване на собствени стойности. Задачата за намиране на собствените стойности и собствените вектори на матрици обикновено възниква при решаване на системи от линейни алгебрични уравнения (СЛАУ) и стохастични диференциални уравнения (СДУ). В практиката, за матрици с големи размерности обикновено е необходимо да се оценят само екстремалните собствени стойности, които определят числото на обусловеност и ключови свойства на системата [63, 87, 90]. Такива задачи възникват в обработката на изображения и сигнали, системите за управление, машинното обучение и квантовата механика, където например енергийните нива на квантовите системи се задават от собствените стойности на Хамилтоновия оператор [64, 73]. Наред с това са широко използвани детерминистични методи като Степенния метод, итерацията на коефициента на Релей, Резолвентния Степенен метод и алгоритъма на Ланцош [33, 43].

Монте Карло методът за оценка на доминиращата собствена стойност е предложен от Собол през 1973 г. [92], а за най-малката собствена стойност - от Михайлов през 1987 г. [55]. Значим принос в развитието на Степенния Монте Карло (СМК) метод имат Димов и Караиванова в периода 1996–2000 г. [15, 18, 21, 23, 24, 25]. През следващите десетилетия са разработени итеративни МК и Квази-Монте Карло (КМК) алгоритми за намиране на екстремални собствени стойности, базирани на Степенния и Резолвентния Степенен метод [3, 19, 21, 25, 60]. Използването на резолвентната матрица за оценка на най-малката собствена стойност е въведено от Димов и Караиванова през 1998 г. [24].

Първите Квази-Монте Карло методи за намиране на екстремални собствени стойности са предложени от Караиванова и Маскани в периода 2001–2003 г. [58, 59, 60, 61]. В редица изследвания са използвани КМК методи, базирани на редици с малък дискрепанс, включително редици на Собол [74, 75, 76], Холтън [39, 40] и Фор [28, 29, 30], които осигуряват по-равномерно и детерминистично покритие на пробното пространство [2, 3, 10, 11, 44, 48]. В сравнение със стандартните МК подходи, КМК методите се отличават с по-добра гъвкавост, приложимост към матрици с големи размерности, относителна простота и висока паралелна ефективност [16, 22, 61].

През последните две десетилетия са разработени и анализирани различни стратегии за паралелизация на Монте Карло алгоритми [4, 5, 6, 78, 79, 83]. Паралелни МК методи за решаване на СЛАУ, базирани на итерацията на Якоби, са представени от Димов [16], както и подходи с минимални вериги на Марков и минимални комуникации [27, 82]. Основна идея в тези разработки е балансирането на стохастичната и систематичната грешка [16]. Разработени са и хибридни Монте Карло алгоритми за матрични пресмятания [1, 84], включително паралелни реализации за оценка на най-голямата и най-малката собствена стойност на матрици, предложени от Димов, Маскани и Караиванова [20, 24, 59].

Цели и задачи на дисертация

Основната цел на дисертационния труд е да се предложат и изследват стохастични числени методи за оценяване на собствени стойности, като се разработят и анализират Монте Карло и рандомизирани Квази-Монте Карло алгоритми за приближено пресмятане на екстремните собствени стойности на симетрични квадратни матрици. За постигане на основната цел са формулирани следните задачи:

1. Да се разгледат и изучат Степенен (Резолвентен) Монте Карло метод и рандомизиран Степенен (Резолвентен) Квази-Монте Карло метод за оценяване на екстремалните собствени стойности на симетрични матрици, както и да се анализира баланса между възникващите грешки - стохастична и систематична.
2. Да се разработят и изследват ефективни Монте Карло и рандомизирани Квази-Монте Карло алгоритми, базирани на тези стохастични методи за числено пресмятане на екстремалните собствени стойности на симетрични квадратни матрици, като се използват генератори на псевдослучайни числа и генератори на редици с малък дискрепанс.
3. Да се проведат числени експерименти с разработените алгоритми, като се обоснове изборът на специализиран софтуер за генериране на редици с малък дискрепанс, в частност редици на Собол и Холтън, и числено да се изследва баланса между стохастичната и систематичната грешка в зависимост от определящите ги параметри.
4. Да се формулира и изследва задача от областта на финансовата математика, свързана с управлението на пазарния риск на портфейл от активи, при който оценката на риска се основава на определянето на максималната собствена стойност на корелационна матрица. Да се приложат разработените в Глава 1 Степенен Монте Карло и рандомизиран Степенен Квази-Монте Карло алгоритми за числено оценяване на най-голямата собствена стойност и да се анализира тяхната ефективност при оценяване на пазарния риск.

Методология на изследването

Методологията на изследването в дисертационния труд се основава на фундаментални научни резултати от следните области:

- Теория на вероятностите и математическа статистика - вероятностни пространства; случайни величини и техни характеристики като функции на разпределение, функция на плътност, математическо очакване, дисперсия; закон за големите числа; централна гранична теорема; вероятностна грешка,

доверителен интервал, верига на Марков; извадки на случайни величини; статистически тестове за изследване на генератори на псевдослучайни величини.

- Математически анализ - редици и редове, сходимост на редове, критерии за сходимост, систематична грешка, норми и метрични пространства.
- Числени методи, линейна и изчислителна алгебра - итерационни методи за СЛАУ и собствени стойности; матрични норми; оптимизационни методи; алгоритми за генериране на псевдослучайни числа, стохастични числени методи (Монте Карло и Квази-Монте Карло методи).
- Теория на числата и по-точно раздела за дискретна и аналитична теория на числата - равномерно разпределени редици, разбъркване (англ. *discrepancy*), радикално-инверсни функции, разлагане по прости основи.
- Информатика/компютърни науки - алгоритми и структури от данни; програмни кодове, написани на МАТЛАВ, С++ с използване на генератори на псевдослучайни числа като MT и MS; библиотеки за рандомизирани редици на Собол и Холтън.

Част от числените експерименти са извършени на високопроизводителния клъстър, който се състои от 12 сървъра Fujitsu Primergy RX 2540 M4. Всеки сървър е конфигуриран с GPU карта от вида NVIDIA Tesla V100 32GB и CPU 2 x Intel Xeon Gold 5118 2.30GHz 24 cores, всеки с 128 GB RAM. Редиците на Собол са получени чрез BRODA's Sobol Randomized Sequence Generator (RSG) [7].

В следващите секции сме формулирали основната задача за оценяване на собствени стойности на симетрична матрица и сме описали МК и КМК подходите за числено решаване на многомерни задачи и грешките, с които се характеризират тези методи.

Формулиране на задачата

Нека разгледаме матрицата $A = \{a_{ij}\}_{i,j=1}^n \in \mathbb{R}^{n \times n}$, $a_{ij} = a_{ji}$, $i, j = 1, \dots, n$. Задачата за намиране на собствени стойности $\lambda(A)$ се формулира по следния начин:

$$A\mathbf{x} = \lambda\mathbf{x}, \quad \mathbf{x} \in \mathbb{R}^n \text{ е произволен ненулев вектор.} \quad (1)$$

Всички собствени стойности на симетричните матрици са реални числа. Без ограничение на общността предполагаме следната наредба на собствените стойности на A :

$$\lambda_{min} = |\lambda_n| \leq |\lambda_{n-1}| \leq \dots \leq |\lambda_2| < |\lambda_1| = \lambda_{max}. \quad (2)$$

Матрицата A може да се представи векторно и по следния начин:

$$A = \{a_{ij}\}_{i,j=1}^n = (\mathbf{a}_1, \dots, \mathbf{a}_i, \dots, \mathbf{a}_n)^T, \text{ където } \mathbf{a}_i = (a_{i1}, \dots, a_{in}), \quad i = 1, \dots, n.$$

Символът T означава операцията *транспониране*. За нуждите на изследването се използват следните векторни и матрични норми:

$$\|\mathbf{h}\| = \|\mathbf{h}\|_1 = \sum_{i=1}^n |h_i|, \quad \|\mathbf{a}_i\| = \|\mathbf{a}_i\|_1 = \sum_{j=1}^n |a_{ij}|, \quad \|A\| = \|A\|_1 = \max_j \sum_{i=1}^n |a_{ij}|, \quad j = 1, \dots, n.$$

Прилагането на стохастични числени методи за оценка на екстремалните собствени стойности λ_{min} и λ_{max} на така дефинираната задача, включва разработването и изследването на Степенни (Резолвентни) Монте Карло и рандомизирани Степенни (Резолвентни) Квази-Монте Карло алгоритми за тяхното приближено пресмятане. Под стохастични числени методи по-нататък ще разбираме Монте Карло методите и тяхната разновидност, рандомизираните Квази-Монте Карло методи.

Монте Карло методи

Монте Карло (МК) методите се основават на конструиране на случайна величина θ , която всъщност представлява измерима функция $\theta = \theta(\omega)$, дефинирана във вероятностното пространство $(\Omega, \mathfrak{A}, \mathbb{P})$ [88], с дефиниционна област Ω и с област на стойностите \mathbb{R} (множеството на реалните числа). Условието при конструирането на случайната величина θ е свързано с математическото ѝ очакване, като изискването е

математическото очакване да съществува и да е равно на търсеното решение на дадена задача или на функционал от решението на изходната задача.

$$E[\theta] = I, \quad (3)$$

където I е търсеното решение (или стойността на функционал от решението). По определение $E[\theta]$ съществува тогава и само тогава, когато съществува $E|\theta|$.

В зависимост от конкретната задача, случайната величина θ може да бъде дискретна или непрекъсната. В дискретния случай математическото очакване се задава чрез:

$$E[\theta(\omega)] = \sum_i \mathbb{P}\{\theta(\omega) = x_i\}x_i = \sum_i p_i x_i = I, \text{ където } \sum_i p_i = 1,$$

а в непрекъснатия случай:

$$E[\theta] = \int_{-\infty}^{\infty} xp(x) dx, \text{ където } \int_{-\infty}^{\infty} p(x) dx = 1,$$

където $p(x) \geq 0$ е вероятностна плътност на разпределение. Функцията на разпределение $F_\theta(x) = \mathbb{P}\{\theta < x\}$ описва натрупаната вероятност и е свързана с $p(x)$ чрез стандартните диференциални зависимости [88, 91]. При реализация на МК алгоритъм се генерира извадка $\theta_1, \theta_2, \dots, \theta_N, \dots$ от независими и еднакво разпределени реализации на случайната величина θ , като приближената стойност на I се определя чрез средната стойност:

$$\bar{\theta}_N = \frac{1}{N} \sum_{i=1}^N \theta_i. \quad (4)$$

Сходимост и оценка на грешката при МК методите

Теоретичната обосновка на МК методите се базира на закона за големите числа (ЗГЧ). Когато математическото очакване е крайно, редицата от средни стойности $\bar{\theta}_N$ (при $N \rightarrow \infty$) е сходяща по вероятност към I (слаб закон на големите числа, теорема на Хинчин) [14, 88]. Необходимото и достатъчно условие, за да бъде сходимостта почти сигурна (усилен закон на големите числа, теорема на Колмогоров), е да има математическо очакване.

Предполагаме, че случайната величина θ има крайна дисперсия:

$$Var[\theta] = E[\theta - E\theta]^2 = E[\theta^2] - (E[\theta])^2 < \infty.$$

Това централната гранична теорема (ЦГТ) осигурява асимптотично нормално разпределение на грешката [31, 88]. В резултат се получава вероятностната оценка:

$$\lim_{N \rightarrow \infty} \mathbb{P} \left\{ \left| \frac{1}{N} \sum_{i=1}^N (\theta_i - I) \right| < x\sigma N^{-1/2} \right\} = \Phi(x),$$

където $\sigma = \sqrt{Var[\theta]}$, а $\Phi(x) = \frac{2}{\sqrt{2\pi}} \int_0^x e^{-t^2/2} dt$ се нарича вероятностен интеграл и се определя от функцията на разпределение на стандартна нормална случайна величина в интервала $(0, 1)$ [92].

Величината

$$r_N = x_\beta \sigma N^{-\frac{1}{2}} \quad (5)$$

се нарича вероятностна грешка, като параметърът β определя нивото на доверие. При $\beta = 0.5$ се получава т.нар. вероятна грешка, която характеризира типичната скорост на сходимост на МК методите от порядъка $\mathcal{O}(N^{-1/2})$.

Дисперсията на случайната величина θ в практическите приложения обикновено се оценява емпирично чрез формулата:

$$Var[\theta] \approx \frac{1}{N} \sum_{i=1}^N (\theta_i)^2 - (\bar{\theta}_N)^2. \quad (6)$$

При неголяма размерност на извадката N от стойности на случайната величина, обикновено се използва

по-точната формула [92]:

$$\text{Var}[\theta] \approx \frac{1}{(N-1)} \sum_{i=1}^N (\theta_i)^2 - \frac{N}{N-1} (\bar{\theta}_N)^2. \quad (7)$$

Разграничават се два типа МК методи, които се използват за намаляване на грешката [22, 89, 92]:

- Ефективни МК методи, при които се цели да се минимизира дисперсията чрез техники за редуциране (отделяне на главна част, симетризация на подинтегралната функция, метод на съществената извадка, метод на разделяне по важност и други), без да се променя асимптотичният порядък на грешката.
- Суперсходящи МК методи, при които чрез симулиране на стойностите на случайна величина в равномерни непокриващи се подобласти и при допълнителни условия за гладкост на подинтегралната функция се постига по-висок порядък на сходимост $r_N = \mathcal{O}(N^{-1/2-\varepsilon(k)})$, където $\varepsilon(k) > 0$ зависи от размерността на задачата.

За симулиране на стойностите на случайните величини се използват генератори на псевдослучайни числа.

Генератори на псевдослучайни числа

Практическата реализация на МК методите изисква използването на генератори на псевдослучайни числа (PRNGs), които генерират равномерно разпределени стойности в $(0, 1)$ [53, 81]. Основните изисквания към добрите PRNGs включват статистическа случайност, дълъг период на повтаряемост, възпроизводимост, ефективност и възможност за паралелно генериране [50, 53].

Сред най-често използваните генератори са линейните и комбинираните конгруентни генератори [32, 54, 71], генераторите с изместени регистри (например, Mersenne Twister [62]), както и генератори на базата на други математически конструкции (например, Middle Square PRNG [70]). В дисертационните изследвания за реализацията на стохастичните алгоритми за оценяване на екстремални собствени стойности са използвани Mersenne Twister (MT) и Middle Square (MS) генератори на псевдослучайни числа.

Квази-Монте Карло методи

Квази-Монте Карло (КМК) методите са детерминистична алтернатива на МК методите и използват редици с малък дискрепанс вместо псевдослучайни точки [39, 66, 74]. Те осигуряват по-равномерно покритие на единичния хиперкуб $[0, 1]^k$ и водят до по-добри скорости на сходимост от порядък $\mathcal{O}(N^{-1}(\log N)^k)$ при достатъчна гладкост на подинтегралната функция [9, 52]. При оценяване на многомерни интеграли приближеното решение има вида:

$$\int_{[0,1]^k} f(x) dx \approx \frac{1}{N} \sum_{i=1}^N f(\mathbf{x}_i), \quad (8)$$

където x_i са елементи на редица с малък дискрепанс.

Редици с малък дискрепанс и оценка на грешката

Ключовата характеристика на тези редици е дискрепанс-звездата $D_N^*(S)$, която измерва отклонението от равномерно разпределение:

$$D_N^*(S) = \sup_{\mathbf{z} \in [0,1]^k} \left| \frac{1}{N} \sum_{i=1}^N \mathbf{1}_Z(\mathbf{x}_i) - \prod_{j=1}^k z_j \right|. \quad (9)$$

Редицата е с малък дискрепанс, ако

$$D_N^*(S) = \mathcal{O}(N^{-1}(\log N)^k) \text{ или } D_N^*(S) \leq C_k \frac{(\log N)^k}{N}, \quad (10)$$

където константата C_k в (10) зависи само от размерността k .

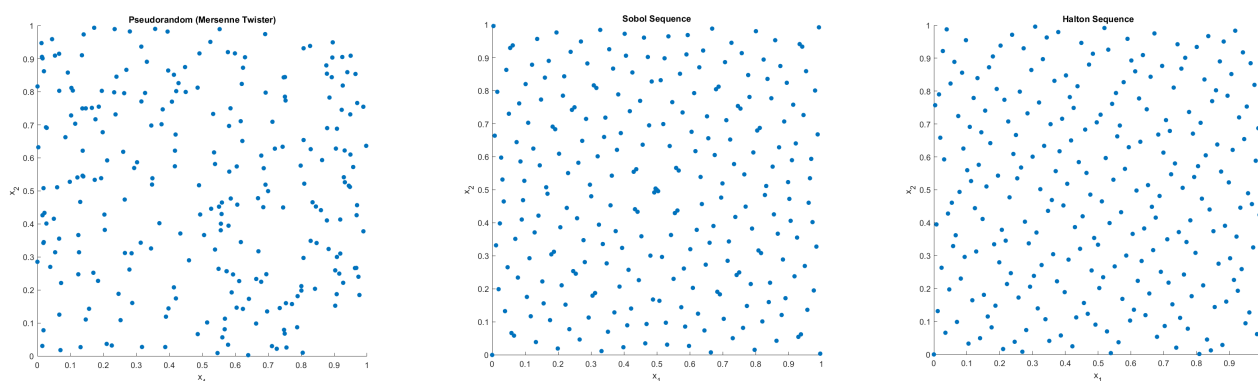
Грешката при КМК методите се оценява чрез неравенството на Коксма-Хлавка [9, 72]:

$$\left| I(f) - \frac{1}{N} \sum_{i=1}^N f(\mathbf{x}_i) \right| \leq V(f) D_N^*(S), \quad (11)$$

където $D_N^*(S)$ е дискрепанс-звезда на редицата от точки $S = \{\mathbf{x}_i\}_{i=1}^{\infty} \in [0, 1]^k$. $V(f)$ е вариацията на функцията в смисъл на Харди-Краус. Това неравенство показва, че грешката зависи както от гладкостта на подинтегралната функция, така и от равномерността на използваната точкова редица. При достатъчно гладки функции се достига асимптотичен порядък на грешката $\mathcal{O}(N^{-1}(\log N)^k)$, който е теоретично по-добър от този на класическите МК методи.

Първите конструкции на редици с малък дискрепанс произлизат от редицата на Ван дер Корпут [85]. Холтън разширява тази идея към многомерния случай чрез използване на радикална инверсия в различни прости основи по координати [39]. Редиците на Холтън са концептуално прости, но при по-високи размерности могат да проявят корелации, което налага използването на модификации [41, 86]. Редиците на Соболев, въведени през 1976 г. [76], представляват цифрови (t, k) -редици при основа 2. Те се конструират чрез направляващи числа, получени от примитивни полиноми над крайното поле $\mathbb{F}_2 = \{0, 1\}$ (събиране по модул 2), което осигурява висока равномерност и добри проекционни свойства [74, 89]. Поради тези свойства редиците на Соболев са сред най-широко използваните в КМК приложения.

Фигура (1) представя сравнение между двумерна псевдослучайна редица, генерирана с Mersenne Twister (MT) и двумерни проекции на 16-мерни редици на Соболев и Холтън. Наблюдава се ясно по-равномерно покритие на единичния квадрат при редиците с малък дискрепанс в сравнение с псевдослучайната редица.



(а) MT генератор на псевдослучайни числа

(б) редица на Соболев

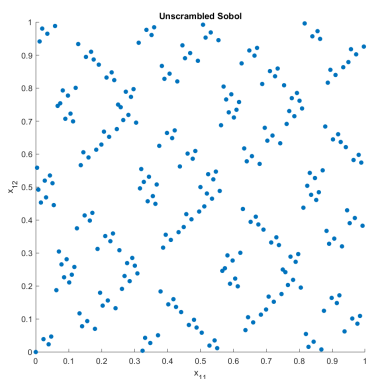
(в) редица на Холтън

Фигура 1: Сравнение на двумерна псевдослучайна редица, получена с MT PRNG, с двумерни проекции на 16-мерни редици на Соболев и Холтън. Генерираните точки са 256, а проекцията е направена по първите две координати.

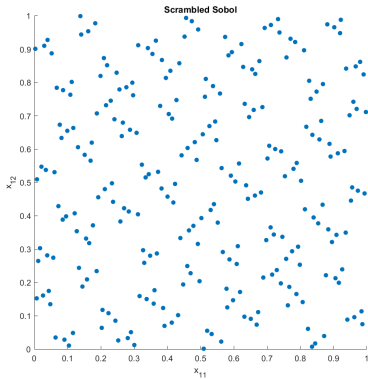
Рандомизирани редици с малък дискрепанс

Детерминистичният характер на класическите редици с малък дискрепанс води до липса на естествена вероятностна рамка за оценка на грешката и до риск от корелации при високи размерности. Тези недостатъци се премахват чрез въвеждане на рандомизирани редици с малък дискрепанс, при които чрез контролирана стохастичност се запазва равномерното покритие на единичния хиперкуб, като едновременно с това се преодоляват ограниченията на фиксираната геометрия [13]. Допълнително към разбърканите редици с малък дискрепанс могат да се прилагат техники за пропускане на начални членове на редицата, чрез използване на оператор за прескачане (англ. *skipping operator*) и оператор за прескачане през стъпка (англ. *leaping operator*) на редицата. Така се запазват свойствата на редиците с малък дискрепанс, като едновременно се въвежда контролирана стохастичност, позволяваща статистическа оценка на грешката. Фигура (2) и Фигура (3) показват наличието на корелации при неразбъркана редица на Соболев и Холтън в по-високи координати и ефекта от прилагането на метода на разбъркване (англ. *scrambling*). Разбъркването води до значително подобряване на равномерността на точковото разпределение, както е показано на Фигура (2б) и Фигура (3б). На Фигура (2в) и Фигура (3в) е показан и допълнителният ефект от комбинирание на метода на разбъркване чрез параметрите *skip* и *leap*. Тези техники редуцират корелациите и подобряват практическите свойства на редиците при задачи с по-висока размерност.

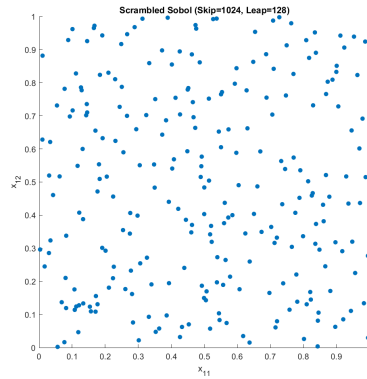
Съвременните модификации на редиците на Холтън и Соболев включва рандомизиране, целящо намаляване на корелациите и подобряване на сходимостта [52, 66] за конкретни задачи. Рандомизацията се реализира чрез прилагане на специално конструирани трансформации (разбърквания) върху вече изградени редици с малък дискрепанс. Съществуват два основни подхода за разбъркване (англ. *scrambling*). Първият се основава на *побитово (цифрово) случайно пермутиране*, при което цифрите на координатите в основа b



(а) неразбъркана редица на Собол

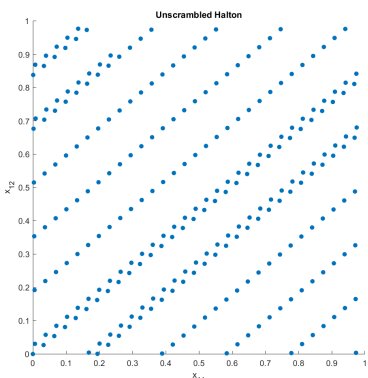


(б) разбъркана редица на Собол

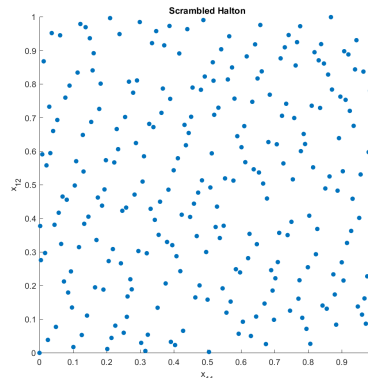


(в) разбъркана редица на Собол с параметри $skip$, $leap$

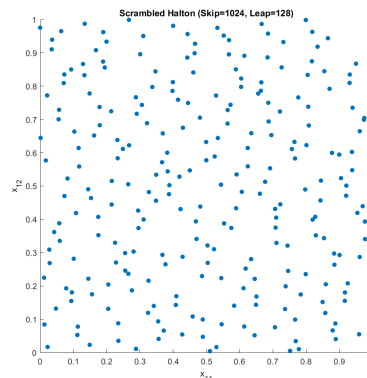
Фигура 2: Сравнение между двумерна проекция с наличие на корелация при 11-та и 12-та координата на редицата на Собол и съответните двумерни проекции на рандомизираните редици на Собол без и с допълнително включени параметри $skip = 1024$ и $leap = 128$.



(а) неразбъркана редица на Холтън



(б) разбъркана редица на Холтън



(в) разбъркана редица на Холтън с параметри $skip$, $leap$

Фигура 3: Сравнение между двумерна проекция с наличие на корелация при 11-та и 12-та координата на редицата на Холтън и съответните двумерни проекции на рандомизираните редици на Холтън без и с допълнително включени параметри $skip = 1024$ и $leap = 128$.

(в частност при основа 2 при двоично представяне) се подлагат на случайни пермутации [67]. Вторият подход се основава на *събиране по модул 1* със случайно число или случаен вектор, генерирани чрез псевдослучаен генератор. Този метод е известен още като *случайно изместване* (англ. *shift*) по модул 1 и се характеризира с проста и ефективна реализация [45, 46]. На практика двата подхода често се комбинират, което води до множество техни варианти и комбинации [68, 69].

Стохастични методи с вериги на Марков

Монте Карло (МК) и Квази-Монте Карло (КМК) методите условно се класифицират в две основни групи: директни и итерационни.

Директните МК/КМК методи се характеризират с това, че полученото приближено решение е повлияно единствено от стохастичната (статистическа) грешка. При класическия МК подход тази грешка е от порядък $\mathcal{O}(N^{-1/2})$ и се оценява чрез неравенство (5), докато при КМК методите тя се описва чрез неравенството на Коксма-Хлавка (11) и има асимптотичен порядък $\mathcal{O}(N^{-1}(\log N)^k)$ при използване на редици с малък дискрепанс, като редиците на Собол и Холтън. В дисертационния труд се прилагат рандомизирани редици с малък дискрепанс, които съчетават по-равномерно покритие на единичния куб с контролирана стохастичност. Поради това грешките, оценявани чрез (5) и (11), се интерпретират като стохастична грешка, зависеща от размера на извадката N .

Итерационните МК/КМК методи, освен стохастична, съдържат и систематична грешка, тъй като оце-

няват не директно точното решение, а крайна итерация на сходящ итерационен процес. Аналогично на класическите числени методи в линейната алгебра, като метода на простата итерация и метода на Гаус–Зайдел [90], систематичната грешка зависи от свойствата на итерационната схема и от нормата на матрицата, участваща в процеса.

В стохастичните итерационни методи за задачи от линейната алгебра се конструира случайна величина $\theta^{(k)}$ чрез верига на Марков с k прехода [16, 22, 92]. Броят на преходите може да бъде фиксиран или случаен с математическо очакване, равно на k . Математическото очакване на случайната величина $\theta^{(k)}$ апроксимира търсеното решение чрез функционал, включващ матрично-векторни произведения. В този контекст систематичната грешка се определя от броя на итерациите k , докато стохастичната грешка зависи от размера на извадката N . Следователно, балансирането между систематичната и стохастичната грешка, както и редукцията на дисперсията, представляват основни изследователски задачи при разработването на ефективни стохастични итерационни методи и алгоритми за линейна алгебра.

Структура на дисертацията

Дисертационният труд се състои от увод, три глави, заключение, приложение и библиография. Дисертацията съдържа 118 страници, включва 15 фигури и 16 таблици, библиография от 121 заглавия, както и списък с математически означения и използвани абревиатури.

- В **Увода** е представена актуалността на тематиката, направен е обзор на основните научни резултати в областта, поставени са целите и задачите на дисертационния труд и подробно е описана методологията на изследването.
- **Глава 1** „Стохастичен Степенен метод: разработване и изследване на ефективни алгоритми за оценяване на максимална собствена стойност“ е посветена на разработването на ефективни стохастични степенни алгоритми за намиране на максималната собствена стойност на плътни симетрични матрици. Дефинирана е задачата и са разгледани методите, които се използват за нейното решаване, а именно, детерминистичния Степенен метод и стохастичния Степенен метод, в неговите два варианта - Степенен Монте Карло (СМК) и Степенен Квази-Монте Карло (СКМК) метод. Подробно е описана конструкцията на случайната величина с използване на верига на Марков, като се акцентира на използването на почти оптимални вероятности. Представен е псевдокод на разработените почти оптимални СМК и СКМК алгоритми и са определени условията за балансиране на стохастичната и систематичната грешка. Главата завършва с дискусия върху направените числени експерименти за матрици с различна размерност.
- **Глава 2** „Стохастичен Резолвентен метод: разработване и изследване на ефективни алгоритми за оценяване на минимална собствена стойност“ разглежда задачата за оценяване на минималната собствена стойност на симетрични квадратни матрици, чрез използване на нейната резолвентна матрица. Представен е Резолвентният Степенен метод и неговите стохастични реализации чрез Резолвентен Монте Карло (РМК) и Резолвентен Квази-Монте Карло (РКМК) подход. В главата се изследва ролята на параметрите, влияещи върху сходимостта на безкрайния ред, чрез който се представя резолвентната матрица. Конструирани са почти оптимални РМК и РКМК алгоритми и се изследва балансът на стохастичната и систематична грешка. Този баланс зависи от броя реализации на случайните величини, дължината на веригата на Марков, степента на резолвентната матрица и ускоряващия параметър, включен в представянето на резолвентната матрица като степенен ред. В последната секция са представени проведените числени експерименти, получени при оценяването на минималната собствена стойност на тестови симетрични матрици.
- **Глава 3** „Стохастични методи за оценяване на пазарния риск на инвестиционни портфейли“ е свързана с прилагането на разработените алгоритми от Глава 1 към реална задача от областта на финансовата математика, свързана с оценяване на пазарния риск на инвестиционен портфейл. Оценяването му се извършва чрез максималната собствена стойност на корелационната матрица на портфейла. Намирането на най-голямата собствена стойност се извършва чрез почти оптималните СМК и СКМК алгоритми. При прилагането на почти оптималния СКМК алгоритъм са разгледани два варианта: (1) с разбъркани редици на Собол и Холтън с настройка по „подразбиране“; (2) с разбъркани редици на Собол и Холтън с използване на параметри `skip` и `leap`. Проведени са числени експерименти с реални финансови данни, които демонстрират приложимостта и ефективността на предложените алгоритми.

В заключителната част са включени: списък на публикациите, апробация, научните и научно-приложните приноси и благодарности.

1. Стохастичен Степенен метод: разработване и изследване на ефективни алгоритми за оценяване на максимална собствена стойност

1.1 Постановка на задачата

Нека $A = \{a_{ij}\}_{i,j=1}^n \in \mathbb{R}^{n \times n}$ е дадена несингулярна симетрична матрица. Собствените стойности (спектърът) на матрицата A се дефинират като реалните стойности на параметъра $\lambda(A)$, за които уравнението:

$$A\mathbf{x} = \lambda(A)\mathbf{x}, \quad \text{където } \mathbf{x} \in \mathbb{R}^n, \quad (1.1)$$

има ненулево решение, като съответните ненулеви вектори \mathbf{x} се наричат собствени вектори на A . Необходимо и достатъчно условие за съществуването на ненулево решение е параметърът $\lambda(A)$ да бъде корен на характеристичното уравнение $|A - \lambda(A)\mathbf{I}| = 0$, където \mathbf{I} е единичната n -мерна матрица [90].

Предполага се, че собствените стойности на A са подредени по абсолютна стойност и удовлетворяват неравенствата:

$$\lambda_{min} = |\lambda_n| \leq |\lambda_{n-1}| \leq |\lambda_{n-2}| \leq \dots \leq |\lambda_2| < |\lambda_1| = \lambda_{max}. \quad (1.2)$$

Разглеждаме задачата за разработване и изследване на ефективни алгоритми на базата на Стохастичния Степенен метод за оценяване на λ_{max} , която съгласно (1.2) е единствена.

1.2 Методи за изчисляване на максималната собствена стойност

1.2.1 Степенен метод

Степенният метод представлява класически итеративен числен алгоритъм за апроксимация на максималната собствена стойност λ_{max} на квадратни матрици [33, 43], базиран на коефициента на Рейли [43]:

$$r(\mathbf{x}) = \frac{\mathbf{x}^T A \mathbf{x}}{\mathbf{x}^T \mathbf{x}},$$

Алгоритъмът се реализира чрез инициализация с нормализиран начален вектор, итеративно обновяване на вектора и оценяване на собствената стойност след k итерации чрез:

$$\lambda_{max} \approx \lambda_{max}^{(k)} = r(\mathbf{x}_k),$$

като при условие (1.2) полученото приближение съответства на доминиращата собствена стойност.

Систематична грешка: В случая на симетрична матрица грешката от прекъсване на Степенния метод при представяне на векторите \mathbf{x}^T и \mathbf{x} като линейни комбинации на ортонормиран базис е от порядъка:

$$|\lambda_{max}^{(k)} - \lambda_{max}| = \mathcal{O}\left(\left|\frac{\lambda_2}{\lambda_1}\right|^{2k}\right), \quad (1.3)$$

което показва силна зависимост от спектралното разстояние между първите две собствени стойности [43].

Сходимость: Скоростта на сходимость на итерационния процес при симетрични матрици се определя от коефициента $\left|\frac{\lambda_2}{\lambda_1}\right|^2$. При близки по модул доминиращи собствени стойности, сходимостьта се забавя и може да изисква прилагане на техники за предварителна обработка.

Изчислителна сложност: Изчислителната сложност на Степенния метод е $\mathcal{O}(kn^2)$, където k е броят на итерациите, а n - размерността на матрицата A , което го прави подходящ за големи, но структурирани симетрични задачи.

1.2.2 Стохастичен Степенен метод

Стохастичният Степенен метод, изследван в дисертационния труд, е разработен от Караиванова и Димов [21, 22]. Той представлява стохастична реализация на *класическия* Степенен метод за задача за собствени стойности (1.1) при условие (1.2). Основата на метода е представянето на най-голямата собствена стойност

като граница на коефициента на Рейли:

$$\lambda_{max} = \lim_{k \rightarrow \infty} \frac{(\mathbf{h}, A^k \mathbf{f})}{(\mathbf{h}, A^{k-1} \mathbf{f})}, \quad (1.4)$$

където \mathbf{f} и \mathbf{h} са произволно избрани и нормирани вектори. Ако k е произволно голямо естествено число, тогава

$$\lambda_{max}^{(k)} = \frac{(\mathbf{h}, A^k \mathbf{f})}{(\mathbf{h}, A^{k-1} \mathbf{f})} \quad (1.5)$$

е приближение на λ_{max} с определена систематична грешка.

Стохастичният Степенен метод се реализира чрез конструиране на случайни величини, чиито математически очаквания съвпадат със скаларните произведения в числителя и знаменателя на горното отношение. За оценяване на $(\mathbf{h}, A^k \mathbf{f})$ се дефинира случайна величина $\theta^{(k)}$, базирана на дискретна крайна верига на Марков със състояния:

$$l_0 \rightarrow \dots \rightarrow l_t \rightarrow \dots \rightarrow l_k \quad (1 \leq l_t \leq n, \quad 0 \leq t \leq k), \quad (1.6)$$

с начален вектор на плътността $\mathbf{p} = \{p_i\}_{i=1}^n \in \mathbb{R}^n$ и матрица на преходната плътност $P = \{p_{ij}\}_{i,j=1}^n \in \mathbb{R}^{n \times n}$. Конструкцията включва итеративно изчисляване на теглото W_k :

$$W_0 = 1, \quad W_t = W_{t-1} \frac{a_{l_{t-1} l_t}}{p_{l_{t-1} l_t}}, \quad t = 1, \dots, k, \quad (1.7)$$

а случайната величина $\theta^{(k)}$ се определя чрез:

$$\theta^{(k)} = \frac{h_{l_0}}{p_{l_0}} W_k f_{l_k}. \quad (1.8)$$

Случайната величина $\theta^{(k)}$, дефинираната чрез (1.8), има математическо очакване равно на скаларното произведение $(\mathbf{h}, A^k \mathbf{f})$, т.е. $E[\theta^{(k)}] = (\mathbf{h}, A^k \mathbf{f})$ [17].

Дефиниция 1.2.1. *Началният вектор на плътността $\mathbf{p} = \{p_i\}_{i=1}^n \in \mathbb{R}^n$ се нарича допустим за вектор $\mathbf{h} = \{h_i\}_{i=1}^n \in \mathbb{R}^n$, т.е. $\mathbf{p} \in \mathfrak{P}_{\mathbf{h}}$, ако $p_i > 0$, когато $h_i \neq 0$ и $p_i = 0$, когато $h_i = 0$ за $i = 1, \dots, n$.*

Дефиниция 1.2.2. *Матрицата на преходната плътност $P = \{p_{ij}\}_{i,j=1}^n \in \mathbb{R}^{n \times n}$ се нарича допустима за симетричната матрица $A = \{a_{ij}\}_{i,j=1}^n \in \mathbb{R}^{n \times n}$, т.е. $P \in \mathfrak{P}_A$, ако $p_{ij} > 0$, когато $a_{ij} \neq 0$ и $p_{ij} = 0$, когато $a_{ij} = 0$ за $i, j = 1, \dots, n$.*

Въведените понятия за допустим начален вектор на плътността $\mathbf{p} \in \mathfrak{P}_{\mathbf{h}}$ и допустима матрица на преходната плътност $P \in \mathfrak{P}_A$ гарантират коректното конструиране на веригата на Марков. За извадка от N реализации на случайната величина $\theta^{(k)}$, средната стойност се представя чрез:

$$\bar{\theta}_N^{(k)} = \frac{1}{N} \sum_{s=1}^N (\theta^{(k)})_s, \quad (1.9)$$

където s -тата реализация на случайната величина $\theta^{(k)}$ е означена с $(\theta^{(k)})_s$ за $1 \leq s \leq N$. Тя представлява стохастична апроксимация на $(\mathbf{h}, A^k \mathbf{f})$. При използване на генератори на псевдослучайни числа вероятностната грешка за Монте Карло метода е:

$$R_N^{(k)} = |(\mathbf{h}, A^k \mathbf{f}) - \bar{\theta}_N^{(k)}| < c_1 \sigma(\theta^{(k)}) N^{-1/2}, \quad (1.10)$$

където c_1 е константа, а $\sigma(\theta^{(k)})$ е стандартно отклонение [17]. При използването на редици с малък дискрепанс (Собол, Холтън), $\bar{\theta}_N^{(k)}$ е Квази-Монте Карло апроксимация с грешка:

$$QR_N^{(k)} = |(\mathbf{h}, A^k \mathbf{f}) - \bar{\theta}_N^{(k)}| < c_2 \frac{(\log N)^{k+1}}{N}, \quad (1.11)$$

получена чрез неравенството на Коксма-Хлавка, където $(k+1)$ е размерността на използваните редици.

Теорема 1.2.1. ¹ *Нека се разгледат случайните величини $\theta^{(k)}$ и $\theta^{(k-1)}$ при фиксираните стъпки k и*

¹Доказателството на Теорема 1.2.1 е разписано в дисертационния труд.

$k - 1$ на веригата на Марков (1.6). Тогава:

$$\frac{E[\theta^{(k)}]}{E[\theta^{(k-1)}]} = \lambda_{max}^{(k)}. \quad (1.12)$$

Забележка 1. Стохастичният Степенен метод се нарича Степенен Монте Карло (СМК) метод, когато извадките за двете случайни величини $\theta^{(k)}$ и $\theta^{(k-1)}$ се генерират чрез генератори на случайни числа за конструиране на веригата на Марков (1.6).

По този начин стохастичната грешка на СМК метода е функция на вероятностната грешка, дефинирана от (1.10).

Забележка 2. Когато за конструирането на веригата на Марков при генериране на извадките на $\theta^{(k)}$ и $\theta^{(k-1)}$ се използват разбъркани редици с малък дискрепанс, като например разбърканите редици на Собол [75], Стохастичният Степенен метод се нарича Степенен Квази-Монте Карло (СКМК) метод.

Стохастичната грешка на СКМК метода е функция на грешката, дефинирана чрез (1.11).

Преформулирането на задачата в термините на интегралните уравнения с подходящи избрани ядра и области позволява използването на неравенството на Коксма-Хлавка при анализиране на сходимостта на разглежданите алгоритми. В общия случай, използването на редици с малък дискрепанс (т. нар. квази-редици) подобрява скоростта на сходимост на съответните КМК алгоритми, но не дава възможност за апостериорна оценка на грешката. Разбъркването (рандомизирането) на квазиредицата позволява избягването на този недостатък.

Теоретичното сравнение на грешката при СМК и СКМК метод показва, че и в двата случая грешката е произведение от два множителя. Единият, зависещ от редицата, а другият от функцията или случайната величина. При това, неравенството на Коксма-Хлавка е граница в най-лошия случай, докато грешката при СМК подход има вероятностен характер. Изчислителният опит показва, че вариацията $V(f)$ в неравенството на Коксма-Хлавка обикновено е свърхоценка, докато вторият множител дискрепансът на редицата показва действителния порядък на грешката.

1.3 Почти оптимални СМК и СКМК методи и алгоритми

1.3.1 Прецизиране на конструкцията на $\theta^{(k)}$ и $\theta^{(k-1)}$ с почти оптимални вероятности за веригата на Марков: балансиране на грешките

Нека симетричната матрица A се представя чрез редовете си \mathbf{a}_i , като се въвеждат стандартни \mathbb{L}_1 -норми за вектори и матрици:

$$\|\mathbf{h}\| = \|\mathbf{h}\|_1 = \sum_{i=1}^n |h_i|, \quad \|\mathbf{a}_i\| = \|\mathbf{a}_i\|_1 = \sum_{j=1}^n |a_{ij}|, \quad \|A\| = \|A\|_1 = \max_j \sum_{i=1}^n |a_{ij}|, \quad j = 1, \dots, n.$$

Вероятностната грешка на МК апроксимациите, съгласно (1.10), зависи от стандартното отклонение $\sigma(\theta^{(k)})$ и броя реализации N . При фиксирано N тя се определя основно от дисперсията на случайната величина $\theta^{(k)}$. Дисперсията зависи от избора на допустимия начален вектор на плътност $\mathbf{p} \in \mathfrak{P}_h$ и матрицата на преходната плътност $P \in \mathfrak{P}_A$, като всички такива избори могат да се използват за конструиране на оценки за $\lambda_{max}^{(k)}$, при условие че дисперсиите са крайни.

Най-често използваният подход за конструиране на верига на Марков за задачи на линейната алгебра се основава на равномерното разпределение. В този случай координатите на вектора на плътност \mathbf{p} и елементите на матрицата на прехода P се задават равни на $\frac{1}{n}$, т.е.:

$$\mathbf{p} = \{p_i\}_{i=1}^n = \frac{1}{n} \text{ и } P = \{p_{ij}\}_{i,j=1}^n = \frac{1}{n}, \quad i, j = 1, \dots, n. \quad (1.13)$$

Когато изборът (1.13) се използва при конструирането на веригата на Марков, Степенният МК и Степенният КМК методи се наричат *класически*. Алтернативни избори на \mathbf{p} и P позволяват конструиране на по-ефективни алгоритми, водещи до редуциране на дисперсията, намаляване на изчислителната сложност и повишаване на точността спрямо *класическия* случай. Съществува *почти оптимален* избор, представящ се чрез формулите:

$$\mathbf{p} = \{p_i\}_{i=1}^n, \quad p_i = \frac{|h_i|}{\|\mathbf{h}\|} \text{ и } P = \{p_{ij}\}_{i,j=1}^n, \quad p_{ij} = \frac{|a_{ij}|}{\|\mathbf{a}_i\|}, \quad i, j = 1, \dots, n. \quad (1.14)$$

който води до значително редуциране на дисперсията, като при този избор конструкцията на случайната величина $\theta^{(k)}$ се различава съществено от *класическия* метод.

Теорема 1.3.1.² Нека \mathbf{p} и P са почти оптималният начален вектор на плътност и почти оптималната матрица на плътност (виж (1.14)). Тогава средната стойност (1.9) може да се представи като:

$$\bar{\theta}_N^{(k)} = \frac{1}{N} \sum_{s=1}^N \text{sign}(h_{l_0^{(s)}}) \|\mathbf{h}\| \left\{ \prod_{t=1}^{t=k} \text{sign}(a_{l_{t-1}^{(s)} l_t^{(s)}}) \| \mathbf{a}_{l_{t-1}^{(s)}} \| \right\} f_{l_k^{(s)}}. \quad (1.15)$$

От Теорема 1.2.1 и 1.3.1 следват следните следствия:

Следствие 1. Нека \mathbf{p} и P са почти оптималният начален вектор на плътност и почти оптималната матрица на плътност. Тогава Стохастичният Степенен метод за оценяване на $\lambda_{max}^{(k)}$ се дефинира като съотношение на средните стойности на случайните величини $\bar{\theta}_N^{(k)}$ и $\bar{\theta}_N^{(k-1)}$ по формула (1.15), т.е.

$$\lambda_{max}^{(k)} \approx \frac{\sum_{s=1}^N \text{sign}(h_{l_0^{(s)}}) \left\{ \prod_{t=1}^{t=k} \text{sign}(a_{l_{t-1}^{(s)} l_t^{(s)}}) \| \mathbf{a}_{l_{t-1}^{(s)}} \| \right\} f_{l_k^{(s)}}}{\sum_{s=1}^N \text{sign}(h_{l_0^{(s)}}) \left\{ \prod_{t=1}^{t=k-1} \text{sign}(a_{l_{t-1}^{(s)} l_t^{(s)}}) \| \mathbf{a}_{l_{t-1}^{(s)}} \| \right\} f_{l_{k-1}^{(s)}}}. \quad (1.16)$$

Следствие 2. Да предположим, че L_1 -нормата на всички вектори-редове на симетричната матрица A е равна на константата a , т.е.

$$\|\mathbf{a}_i\|_1 = \|\mathbf{a}_i\| = \sum_{j=1}^n |a_{ij}| = a, \quad i = 1, \dots, n. \quad (1.17)$$

Тогава оценката за $\lambda_{max}^{(k)}$ по формула (1.16) може да се пренапише по по-прост начин:

$$\lambda_{max}^{(k)} \approx \frac{\sum_{s=1}^N \text{sign}(h_{l_0^{(s)}}) \left\{ \prod_{t=1}^{t=k} \text{sign}(a_{l_{t-1}^{(s)} l_t^{(s)}}) \right\} a f_{l_k^{(s)}}}{\sum_{s=1}^N \text{sign}(h_{l_0^{(s)}}) \left\{ \prod_{t=1}^{t=k-1} \text{sign}(a_{l_{t-1}^{(s)} l_t^{(s)}}) \right\} f_{l_{k-1}^{(s)}}}. \quad (1.18)$$

Следствие 3. Нека L_1 -нормата на всички вектори-редове на симетричната матрица A удовлетворява условието (1.17). Предполага се, че следното условие е валидно за всички елементи на A :

$$\text{sign}(h_{l_0^{(s)}}) \prod_{t=1}^{t=k-1} \text{sign}(a_{l_{t-1}^{(s)} l_t^{(s)}}) = 1, \quad s = 1, \dots, N. \quad (1.19)$$

Тогава $\lambda_{max}^{(k)}$ може да се оцени по по-проста формула, както следва:

$$\lambda_{max}^{(k)} \approx \frac{\sum_{s=1}^N \text{sign}(a_{l_{k-1}^{(s)} l_k^{(s)}}) a f_{l_k^{(s)}}}{\sum_{s=1}^N f_{l_{k-1}^{(s)}}}. \quad (1.20)$$

Чрез Следствие (2) и Следствие (3) се конструират опростени оценки за приближено пресмятане на $\lambda_{max}^{(k)}$ за определени класове симетрични матрици, като съгласно Забележка 3 тези класове не са празно множество.

Забележка 3. Множеството от симетрични матрици, които удовлетворяват условията на Следствие 3, е непразно. Например, нека се разгледа симетрична матрица A с положителни елементи, за която всички елементи по главния диагонал са равни, т.е. $a_{ii} = b > 0$, $i = 1, \dots, n$, и всички извън диагонални елементи също са равни, т.е. $a_{ij} = d > 0$, $i \neq j$, $i, j = 1, \dots, n$.

Тъй като систематичната грешка ϵ_1 зависи от броя k на преходите във веригата на Марков, а стохастичната грешка ϵ_2 зависи от броя на реализациите N , се получава оценка за съответните грешки:

$$\left| \lambda_{max} - \frac{\bar{\theta}_N^{(k)}}{\bar{\theta}_N^{(k-1)}} \right| = \left| \lambda_{max} - \lambda_{max}^{(k)} + \lambda_{max}^{(k)} - \frac{\bar{\theta}_N^{(k)}}{\bar{\theta}_N^{(k-1)}} \right| \leq \left| \lambda_{max} - \lambda_{max}^{(k)} \right| + \left| \lambda_{max}^{(k)} - \frac{\bar{\theta}_N^{(k)}}{\bar{\theta}_N^{(k-1)}} \right| < \epsilon_1 + \epsilon_2, \quad (1.21)$$

²Доказателството на Теорема 1.3.1 е разписано в дисертационния труд.

като оптимален баланс между двата вида грешки (систематичната и стохастичната грешка) се постига при условие $\epsilon_1 = \epsilon_2 = \epsilon$.

Дефиниция 1.3.1. *СМК и СКМК методи за оценка на максималната собствена стойност на симетрична матрица при избор на начален вектор на плътност \mathbf{p} и матрица на плътност на прехода P съгласно формула (1.14) се наричат почти оптимални СМК и СКМК методи.*

1.3.2 Почти оптимален СМК и почти оптимален СКМК алгоритми: изчислителна сложност

Псевдокодът в **Алгоритъм 1** описва процедура за оценяване на максималната собствена стойност на симетрична плътна матрица въз основа на формула (1.16) и Следствие 1. Алгоритъмът се определя като *почти оптимален* СКМК алгоритъм при използване на редици с малък дискрепанс (Собол или Холтгън) в стъпка 4, а при прилагане на генератори на псевдослучайни числа - като *почти оптимален* СМК алгоритъм, означаван с **Алгоритъм 1а**. При изпълнение на условията от Следствие 2 и Следствие 3 формулите в стъпка 6 могат да бъдат заменени с по-опростените изрази (1.18) или (1.20).

Algorithm 1 Pseudocode for computing the *Almost Optimal* PQMC algorithm

- 1: INPUT: matrix $A = \{a_{ij}\} \in \mathbb{R}^{n \times n}$; vectors $\mathbf{h}, \mathbf{f} \in \mathbb{R}^n$ and positive integers n, N, k
- 2: COMPUTE: $\|\mathbf{a}_i\| = \sum_{j=1}^n |a_{ij}|$, $1 \leq i \leq n$ and $\|\mathbf{h}\| = \sum_{i=1}^n |h_i|$
- 3: COMPUTE: The coordinates of a permissible density vector \mathbf{p} and the elements of a permissible density matrix P

$$p_i = \frac{|h_i|}{\|\mathbf{h}\|} \quad \text{and} \quad p_{ij} = \frac{|a_{ij}|}{\|\mathbf{a}_i\|}, \quad i, j = 1, \dots, n.$$

- 4: GENERATE: N elements of the $(k+1)$ -dimensional Sobol (or Halton) sequence
- 5: CONSTRUCT: N realizations of the Markov chain with integer elements

$$l_0^{(s)} \rightarrow \dots \rightarrow l_t^{(s)} \rightarrow \dots \rightarrow l_{k-1}^{(s)} \rightarrow l_k^{(s)}, \quad 1 \leq l_t^{(s)} \leq n, \quad 0 \leq t \leq k, \quad 1 \leq s \leq N$$

- 6: COMPUTE:

$$\bar{\theta}_N^{(k)} = \frac{1}{N} \sum_{s=1}^N \text{sign}(h_{l_0^{(s)}}) \|\mathbf{h}\| \left\{ \prod_{t=1}^{t=k} \text{sign}(a_{l_{t-1}^{(s)} l_t^{(s)}}) \|\mathbf{a}_{l_{t-1}^{(s)}}\| \right\} f_{l_k^{(s)}};$$

$$\bar{\theta}_N^{(k-1)} = \frac{1}{N} \sum_{s=1}^N \text{sign}(h_{l_0^{(s)}}) \|\mathbf{h}\| \left\{ \prod_{t=1}^{t=k-1} \text{sign}(a_{l_{t-1}^{(s)} l_t^{(s)}}) \|\mathbf{a}_{l_{t-1}^{(s)}}\| \right\} f_{l_{k-1}^{(s)}};$$

- 7: OUTPUT:

$$\lambda_{max}^{(k)} := \frac{\bar{\theta}_N^{(k)}}{\bar{\theta}_N^{(k-1)}}$$

Изчислителната сложност на **Алгоритъм 1** се определя от броя реализации N на веригата на Марков, средното време за изчисляване на една реализация на случайните величини $\theta^{(k)}$ и $\theta^{(k-1)}$, както и от предварителното конструиране на матрицата на преходната плътност P . Стъпки 2 и 3 изискват $\mathcal{O}(n^2)$ операции и се разглеждат като предварителна обработка, докато основната изчислителна тежест е съсредоточена в стъпки 4–6. За фиксирано k времето за една реализация е $\tau_k = \mathcal{O}(nk)$, поради което общата изчислителна сложност на алгоритъма е:

$$\text{Cost}(\text{Alg1}) = \mathcal{O}(Nnk) + \mathcal{O}(n^2) = \mathcal{O}(Nnk + n^2),$$

която при $N > n$ се свежда до $\mathcal{O}(Nnk)$. Числените експерименти показват, че за плътни симетрични матрици балансът между систематичната и стохастичната грешка се постига при $N = \mathcal{O}(n)$, като при разреждени матрици може да се очаква допълнително намаляване на изчислителните разходи.

1.4 Числени резултати и дискусия

В числените експерименти **Алгоритъм 1а** е реализиран с два генератора на псевдослучайни числа: Mersenne Twister (MT) и Middle Square (MS), като съответните реализации са означени като $PMC_{(MT)}$ и $PMC_{(MS)}$. Получените резултати са сравнени с тези на *класическия* СМК алгоритъм, реализиран чрез

стандартния избор на началния вектор и матрицата на преходната плътност по формула (1.14), означени като $PMC_{(MT)}^{(n)}$ и $PMC_{(MS)}^{(n)}$. **Алгоритъм 1** е реализиран с редици с малък дискрепанс - Соболев и Холтън, означени като $PQMC_{(S)}$ и $PQMC_{(H)}$, и е сравнен с *класическия* СКМК подход и неговите варианти $PQMC_{(S)}^{(n)}$ и $PQMC_{(H)}^{(n)}$. За намаляване на корелациите при големи размерности, редиците са подложени на разбъркване: "RR2" (*reverse-radix*) за Холтън и "MatousekAffineOwen" за Соболев. Всички генератори и редици са внедрени в MATLAB 9.10 (R2021a).

Числените експерименти са проведени върху две симетрични плътни матрици с размерности $n = 100$ и $n = 500$, със случайно генерирани елементи и със собствени стойности съответно $(\lambda_1, \lambda_2) = (50.0408, 4.0522)$ и $(\lambda_1, \lambda_2) = (250.2454, 9.0721)$. Векторите \mathbf{h} и \mathbf{f} са избрани с L_1 -норма. За да се осигури сходимостта, матрицата A е мащабирана с параметъра:

$$q = \frac{1}{|\lambda_1| + \delta_1}, \quad \delta_1 > 0,$$

или алтернативно:

$$q = \frac{1}{\max_i \|\mathbf{a}_i\| + \delta}, \quad \delta > 0, \quad i = 1, \dots, n,$$

където $\|\mathbf{a}_i\|$ е L_1 -нормата на i -тия ред на A . Подборът на добре разделени собствени стойности гарантира бърза сходимост и позволява експериментално изследване на баланса между систематичната и стохастичната грешка. Анализирано е влиянието на броя преходи k и броя реализации N върху точността, както и ефектът от редуцираната дисперсия при *почти оптималните* СКМК/СКМК алгоритми при фиксирана изчислителна сложност.

1.4.1 Случай (i): Размер на матрицата $n = 100$

В Таблица 1.1 са представени резултатите за оценяване на най-голямата собствена стойност на симетрична матрица ($n = 100$, $\lambda_1 = 50.0408$) с *почти оптималния* $PMC_{(MT)}$ и *класическия* $PMC_{(MT)}^{(n)}$ алгоритъм. Оптималният баланс между систематичната и стохастичната грешка се достига при $k = 8$ или $k = 9$, като минималната абсолютна грешка за $PMC_{(MT)}$ е 5×10^{-5} при $k = 8$, $N = 512$, докато за *класическия* $PMC_{(MT)}^{(n)}$ се наблюдава грешка 1.5×10^{-3} при $k = 8$, $N = 2048$, което показва значително превъзходство на *почти оптималния* вариант.

Таблица 1.1: Резултати от $PMC_{(MT)}$ и $PMC_{(MT)}^{(n)}$ при различни N и k за матрица A с размерност $n = 100$, за която $\lambda_2/\lambda_1 \approx 0.081$, където $\lambda_1 = 50.0408$ и $\lambda_2 = 4.0522$.

| N | k | $\lambda_1^{(k)}$ using $PMC_{(MT)}$ | Absolute Error $PMC_{(MT)}$ | $\lambda_1^{(k)}$ using $PMC_{(MT)}^{(n)}$ | Absolute Error $PMC_{(MT)}^{(n)}$ |
|------|----------|---|--------------------------------|---|--------------------------------------|
| 128 | 8 | 50.039901 | 0.00090 | 50.019883 | 0.0209 |
| | 9 | 50.040709 | 0.00009 | 50.037795 | 0.0030 |
| | 10 | 50.045295 | 0.00450 | 50.066104 | 0.0253 |
| 256 | 7 | 50.023923 | 0.01688 | 50.012158 | 0.0286 |
| | 8 | 50.040628 | 0.00017 | 50.039528 | 0.0013 |
| | 9 | 50.044896 | 0.00410 | 50.050593 | 0.0098 |
| 512 | 7 | 50.035045 | 0.00576 | 49.996040 | 0.0448 |
| | 8 | 50.040750 | 0.00005 | 50.039264 | 0.0015 |
| | 9 | 50.052027 | 0.01123 | 50.050158 | 0.0094 |
| 1024 | 7 | 50.029828 | 0.01097 | 50.019435 | 0.0214 |
| | 8 | 50.040720 | 0.00008 | 50.039735 | 0.0011 |
| | 9 | 50.045403 | 0.00460 | 50.054351 | 0.0136 |
| 2048 | 7 | 50.036291 | 0.00451 | 50.029563 | 0.0112 |
| | 8 | 50.041086 | 0.00029 | 50.041728 | 0.0009 |
| | 9 | 50.043850 | 0.00305 | 50.057179 | 0.0164 |

*Маркираните редове представляват оптималните k стойности за всяко N .

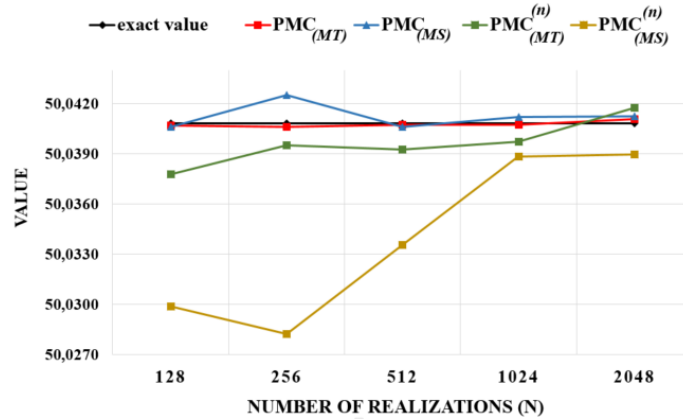
Таблица 1.2 показва аналогичните резултати за $PMC_{(MS)}$ и $PMC_{(MS)}^{(n)}$. Най-малката абсолютна грешка за *почти оптималния* $PMC_{(MS)}$ алгоритъм също се наблюдава при $k = 8$, $N = 512$ и е 1.8×10^{-4} . При *класическия* $PMC_{(MS)}^{(n)}$ алгоритъм тя е 1.8×10^{-3} при $k = 8$, $N = 2048$. Това потвърждава тенденцията, че *почти оптималните* алгоритми редуцират дисперсията и водят до по-точни оценки в сравнение с

класическите, независимо от използвания генератор на псевдослучайни числа.

Таблица 1.2: Резултати от $PMC_{(MS)}$ и $PMC_{(MS)}^{(n)}$ при различни N и k за матрицата A с размерност $n = 100$, за която $\lambda_2/\lambda_1 \approx 0.081$, където $\lambda_1 = 50.0408$ и $\lambda_2 = 4.0522$.

| N | k | $\lambda_1^{(k)}$ using | Absolute Error | $\lambda_1^{(k)}$ using | Absolute Error |
|------|----------|-------------------------|----------------|-------------------------|--------------------|
| | | $PMC_{(MS)}$ | $PMC_{(MS)}$ | $PMC_{(MS)}^{(n)}$ | $PMC_{(MS)}^{(n)}$ |
| 128 | 8 | 50.039440 | 0.00136 | 50.019388 | 0.0214 |
| | 9 | 50.040595 | 0.00020 | 50.029883 | 0.0109 |
| | 10 | 50.051262 | 0.01046 | 50.060923 | 0.0201 |
| 256 | 7 | 50.031044 | 0.00976 | 49.984883 | 0.0559 |
| | 8 | 50.042490 | 0.00169 | 50.028211 | 0.0126 |
| | 9 | 50.050490 | 0.00969 | 50.051251 | 0.0105 |
| 512 | 7 | 50.030518 | 0.01028 | 49.997989 | 0.0428 |
| | 8 | 50.040617 | 0.00018 | 50.033570 | 0.0072 |
| | 9 | 50.058913 | 0.01811 | 50.053570 | 0.0128 |
| | 10 | 50.061048 | 0.02025 | 50.091035 | 0.0502 |
| 1024 | 7 | 50.031472 | 0.00933 | 50.009835 | 0.0310 |
| | 8 | 50.041209 | 0.00041 | 50.038850 | 0.0020 |
| | 9 | 50.044938 | 0.00414 | 50.056396 | 0.0156 |
| 2048 | 7 | 50.038998 | 0.00180 | 50.022658 | 0.0181 |
| | 8 | 50.041243 | 0.00044 | 50.038951 | 0.0018 |
| | 9 | 50.043045 | 0.00225 | 50.063537 | 0.0227 |

*Маркираните редове представляват оптималните k стойности за всяко N .

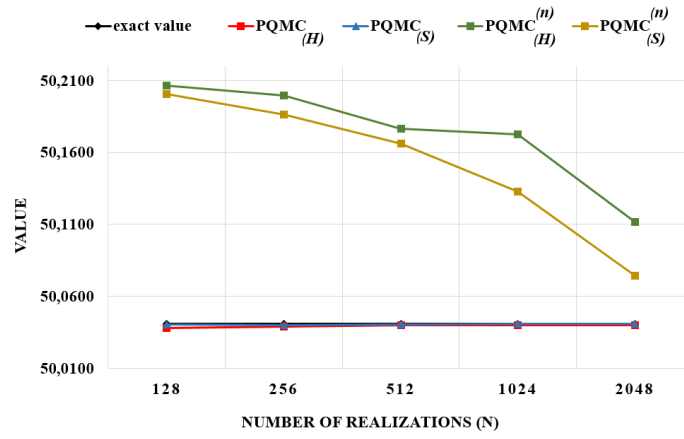


Фигура 1.1: Графично сравняване на приближените резултати за $\lambda_{max}^{(k)}$ с $\lambda_{max} = 50.0408$ при балансиран случай за грешките, т.е. когато $(k = 9, N = 128)$ и $(k = 8, N > 128)$. Приложени алгоритми: $PMC_{(MT)}$, $PMC_{(MS)}$, $PMC_{(MT)}^{(n)}$ и $PMC_{(MS)}^{(n)}$.

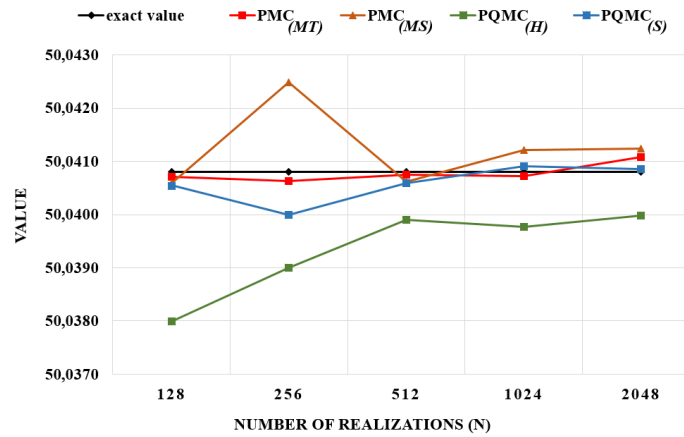
На Фигура 1.1 е показано, че *почти оптималните* версии ($PMC_{(MT)}$ и $PMC_{(MS)}$) следват равномерни траектории близо до точната стойност λ_{max} дори при по-малки N , докато *класическите* варианти, особено $PMC_{(MS)}^{(n)}$, демонстрират по-големи колебания и по-слаба стабилност. Основната причина е намалената дисперсия при конструирането на случайните величини $\theta^{(k)}$ и $\theta^{(k-1)}$ чрез оптималната матрица на преходната плътност, което потвърждава по-високата точност и изчислителна ефективност на *почти оптималните* СКМК алгоритми.

В Таблица 1.3 са представени резултатите за *почти оптималните* и *класическите* СКМК алгоритми, базирани на редиците на Холтън, при оценяване на $\lambda_1 = 50.0408$ за матрица с размер $n = 100$. Наблюдава се монотонно намаляване на абсолютната грешка с увеличаване на броя преходи k , като оптималната сходимост се постига при $k = 11-12$. При $N = 2048$ алгоритъмът $PQMC_{(H)}$ достига абсолютна грешка 8.2×10^{-4} , докато *класическият* $PQMC_{(H)}^{(n)}$ остава значително по-неточен с грешка 0.0707, което ясно демонстрира предимството на *почти оптималната* конструкция.

Таблица 1.4 съдържа аналогичните резултати за алгоритмите, базирани на редиците на Соболев. *Почти оптималният* алгоритъм $PQMC_{(S)}$ систематично превъзхожда както своя *класически* аналог, така и варианта с редици на Холтън. Например, при $N = 2048$ и оптимален брой преходи, $PQMC_{(S)}$ постига



Фигура 1.2: Графично сравняване на приближените резултати за $\lambda_{max}^{(k)}$ с $\lambda_{max} = 50.0408$ при балансирания случай за грешките, т.е. когато $(k = 12, N = 128)$ и $(k = 11, N > 128)$. Приложени алгоритми: $PQMC_{(S)}$, $PQMC_{(H)}$, $PQMC_{(S)}^{(n)}$ и $PQMC_{(H)}^{(n)}$.



Фигура 1.3: Графично сравнение на λ_{max} на симетричната матрица A ($n = 100$) с $\lambda_{max}^{(k)}$ в балансиращите случаи. Използвани са почти оптималните СМК и СКМК алгоритми.

Таблица 1.3: Резултати от $PQMC_{(H)}$ и $PQMC_{(H)}^{(n)}$ при различни N и k за матрица A с размерност $n = 100$, за която $\lambda_2/\lambda_1 \approx 0.081$, където $\lambda_1 = 50.0408$ и $\lambda_2 = 4.0522$.

| N | k | $\lambda_1^{(k)}$ using | Absolute Error | $\lambda_1^{(k)}$ using | Absolute Error |
|------|-----------|-------------------------|----------------|-------------------------|--------------------|
| | | $PQMC_{(H)}$ | $PQMC_{(H)}$ | $PQMC_{(H)}^{(n)}$ | $PQMC_{(H)}^{(n)}$ |
| 128 | 11 | 49.801798 | 0.23900 | 49.574856 | 0.4659 |
| | 12 | 50.037989 | 0.00281 | 50.206208 | 0.1654 |
| | 13 | 50.152986 | 0.11219 | 50.314121 | 0.2733 |
| 256 | 10 | 50.001834 | 0.03897 | 49.739820 | 0.3010 |
| | 11 | 50.038996 | 0.00180 | 50.199525 | 0.1587 |
| | 12 | 50.067986 | 0.02719 | 50.229280 | 0.1885 |
| 512 | 10 | 50.017989 | 0.02281 | 49.749688 | 0.2911 |
| | 11 | 50.039899 | 0.00090 | 50.176724 | 0.1359 |
| | 12 | 50.178512 | 0.13771 | 50.196566 | 0.1558 |
| 1024 | 10 | 49.976987 | 0.06381 | 49.806135 | 0.2347 |
| | 11 | 50.039772 | 0.00103 | 50.172345 | 0.1315 |
| | 12 | 50.107985 | 0.06719 | 50.295866 | 0.2551 |
| 2048 | 10 | 49.973344 | 0.06746 | 49.907160 | 0.1336 |
| | 11 | 50.039982 | 0.00082 | 50.111519 | 0.0707 |
| | 12 | 50.183245 | 0.14245 | 50.275169 | 0.2344 |

*Маркираните редове представляват оптималните k стойности за всяко N .

абсолютна грешка 5×10^{-5} , докато $PQMC_{(S)}^{(n)}$ достига 0.0336. Дори при по-малки стойности на N , като $N = 256$ и $k = 11$, разликата е съществена - 8.1×10^{-4} срещу 0.1455 в полза на *почти оптималния* алгоритъм.

Таблица 1.4: Резултати от $PQMC_{(S)}$ и $PQMC_{(S)}^{(n)}$ при различни N и k за матрица A с размерност $n = 100$, за която $\lambda_2/\lambda_1 \approx 0.081$, където $\lambda_1 = 50.0408$ и $\lambda_2 = 4.0522$.

| N | k | $\lambda_1^{(k)}$ using | Absolute Error | $\lambda_1^{(k)}$ using | Absolute Error |
|------|-----------|-------------------------|---------------------|------------------------------------|------------------------------------|
| | | PQMC _(S) | PQMC _(S) | PQMC _(S) ⁽ⁿ⁾ | PQMC _(S) ⁽ⁿ⁾ |
| 128 | 11 | 49.679383 | 0.36142 | 49.624160 | 0.4166 |
| | 12 | 50.040544 | 0.00026 | 50.200481 | 0.1597 |
| | 13 | 50.091116 | 0.05032 | 50.299826 | 0.2590 |
| 256 | 10 | 49.994318 | 0.04648 | 49.742961 | 0.2978 |
| | 11 | 50.039989 | 0.00081 | 50.186264 | 0.1455 |
| | 12 | 50.049407 | 0.00861 | 50.225854 | 0.1851 |
| 512 | 10 | 50.050000 | 0.00920 | 49.866152 | 0.1746 |
| | 11 | 50.040585 | 0.00021 | 50.166400 | 0.1256 |
| | 12 | 50.148019 | 0.10722 | 50.223877 | 0.1831 |
| 1024 | 10 | 49.986946 | 0.05385 | 49.933245 | 0.1076 |
| | 11 | 50.040913 | 0.00011 | 50.132603 | 0.0918 |
| | 12 | 50.087435 | 0.04664 | 50.167252 | 0.1265 |
| 2048 | 10 | 49.972883 | 0.06792 | 49.957605 | 0.0832 |
| | 11 | 50.040854 | 0.00005 | 50.074364 | 0.0336 |
| | 12 | 50.155992 | 0.11519 | 50.186986 | 0.1462 |

*Маркираните редове представляват оптималните k стойности за всяко N .

Визуализацията на Фигура 1.2 потвърждава числените резултати от Таблица 1.3 и Таблица 1.4, като показва по-гладка и стабилна сходимост на алгоритмите $PQMC_{(S)}$ и $PQMC_{(H)}$ към точната стойност λ_{max} за всички стойности на N . За разлика от тях, *класическите* варианти $PQMC_{(S)}^{(n)}$ и $PQMC_{(H)}^{(n)}$ демонстрират по-големи колебания и начално надценяване на собствената стойност, като сходимостта им е значително по-бавна.

На Фигура 1.3 са показани сравненията между *почти оптималните* СМК и СКМК алгоритми за балансираните случаи. При СМК методите $PMC_{(MT)}$ достига минимална грешка 5×10^{-5} при $N = 512$ и $k = 8$, докато $PMC_{(MS)}$ показва по-голяма, но все още малка грешка 1.8×10^{-4} . От страна на СКМК алгоритмите, $PQMC_{(S)}$ е най-точният, с грешка 5×10^{-5} при $N = 2048$ и $k = 11$, докато $PQMC_{(H)}$ достига 8.2×10^{-4} . Тези резултати ясно показват, че използването на разбъркани редици на Собол и *почти оптималната* преходна матрица води до по-малка дисперсия и оптимален баланс между стохастичната и систематичната грешка спрямо *класическите* им варианти.

1.4.2 Случай (ii): Размер на матрицата $n = 500$

В Таблица 1.5 са представени резултатите за оценяване на максималната собствена стойност λ_{max} на симетрична плътна матрица с размерност $n = 500$, получени с *почти оптималния* алгоритъм $PMC_{(MT)}$ и неговия *класически* аналог $PMC_{(MT)}^{(n)}$. Числените данни показват, че при фиксиран брой реализации N увеличаването на броя преходи k води до намаляване на абсолютната грешка, като оптималният баланс между систематичната и стохастичната грешка се постига при $k = 9$. В този балансиран случай *почти оптималният* алгоритъм достига абсолютна грешка от порядъка на 10^{-4} (например, при $N = 512$), докато при *класическия* вариант грешката остава от порядъка на 10^{-2} (при същите параметри), което ясно демонстрира значително по-висока точност. Аналогични изводи следват от Таблица 1.6, където са представени резултатите за *почти оптималния* алгоритъм $PMC_{(MS)}$ и *класическия* $PMC_{(MS)}^{(n)}$. В този случай оптималната сходимост се наблюдава при $k = 9$, като минималната абсолютна грешка за *почти оптималния* алгоритъм е постигната при $N = 128$ и е от порядъка на 10^{-4} . За разлика от него, *класическият* вариант демонстрира значително по-големи грешки от порядъка на 10^{-2} , което потвърждава предимството на оптимизираната конструкция дори при използване на по-слаб псевдослучаен генератор. Визуализацията на Фигура 1.4 потвърждава резултатите от Таблица 1.5 и Таблица 1.6. *Почти оптималните* СМК алгоритми показват по-гладка и стабилна сходимост към λ_{max} при нарастващ брой реализации N . Това се дължи на значително намалената дисперсия при конструирането на случайните величини $\theta^{(k)}$ и $\theta^{(k-1)}$. *Класическият* алгоритъм $PMC_{(MS)}^{(n)}$ демонстрира колебания и дори надценяване на собствената

Таблица 1.5: Резултати от $PMC_{(MT)}$ и $PMC_{(MT)}^{(n)}$ при различни N и k за матрица A с размерност $n = 500$, за която $\lambda_2/\lambda_1 \approx 0.0363$, където $\lambda_1 = 250.2454$ и $\lambda_2 = 9.0721$.

| N | k | $\lambda_1^{(k)}$ using $PMC_{(MT)}$ | Absolute Error $PMC_{(MT)}$ | $\lambda_1^{(k)}$ using $PMC_{(MT)}^{(n)}$ | Absolute Error $PMC_{(MT)}^{(n)}$ |
|------|----------|---|--------------------------------|---|--------------------------------------|
| 128 | 8 | 250.222396 | 0.02300 | 249.492834 | 0.7526 |
| | 9 | 250.232288 | 0.01311 | 250.032970 | 0.2124 |
| | 10 | 250.295027 | 0.04963 | 250.149293 | 0.0961 |
| 256 | 8 | 250.228998 | 0.01640 | 250.124994 | 0.1204 |
| | 9 | 250.250130 | 0.00473 | 250.222275 | 0.0231 |
| | 10 | 250.283657 | 0.03826 | 250.311181 | 0.0658 |
| 512 | 8 | 250.235302 | 0.01011 | 250.115109 | 0.1303 |
| | 9 | 250.245110 | 0.00029 | 250.192305 | 0.0531 |
| | 10 | 250.268302 | 0.02290 | 250.328365 | 0.0830 |
| 1024 | 8 | 250.235306 | 0.01009 | 250.164200 | 0.0812 |
| | 9 | 250.245015 | 0.00038 | 250.205542 | 0.0399 |
| | 10 | 250.269913 | 0.02451 | 250.300328 | 0.0549 |
| 2048 | 8 | 250.226172 | 0.01923 | 250.175939 | 0.0695 |
| | 9 | 250.246938 | 0.00154 | 250.224397 | 0.0210 |
| | 10 | 250.251779 | 0.00638 | 250.295497 | 0.0501 |

*Маркираните редове представляват оптималните k стойности за всяко N .

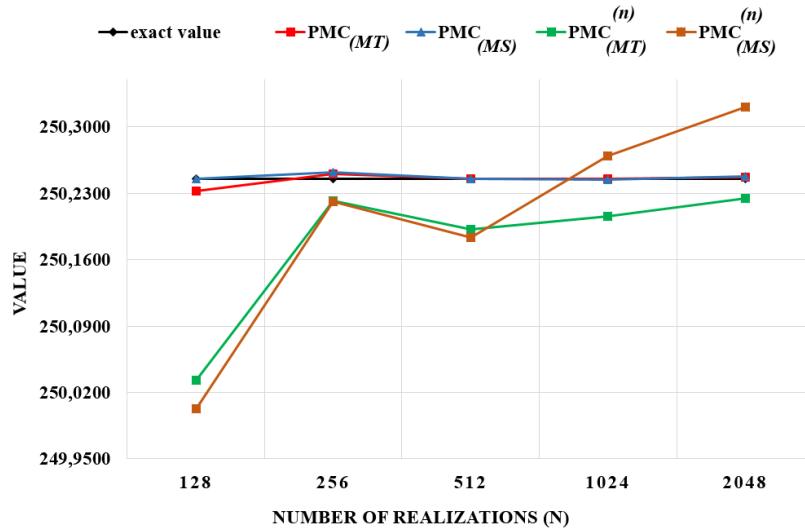
стойност при по-големи стойности на N , докато $PMC_{(MT)}^{(n)}$ се характеризира с по-стабилна сходимост. Това показва, че МТ генераторът на псевдослучайни числа е за предпочитане при задачи с високи размерности.

Таблица 1.6: Резултати от $PMC_{(MS)}$ и $PMC_{(MS)}^{(n)}$ при различни N и k за матрица A с размерност $n = 500$, за която $\lambda_2/\lambda_1 \approx 0.0363$, където $\lambda_1 = 250.2454$ и $\lambda_2 = 9.0721$.

| N | k | $\lambda_1^{(k)}$ using $PMC_{(MS)}$ | Absolute Error $PMC_{(MS)}$ | $\lambda_1^{(k)}$ using $PMC_{(MS)}^{(n)}$ | Absolute Error $PMC_{(MS)}^{(n)}$ |
|------|----------|---|--------------------------------|---|--------------------------------------|
| 128 | 8 | 250.211445 | 0.03395 | 249.462722 | 0.7827 |
| | 9 | 250.245247 | 0.00015 | 250.003068 | 0.2423 |
| | 10 | 250.279476 | 0.03408 | 250.120230 | 0.1252 |
| 256 | 8 | 250.203720 | 0.04168 | 250.015811 | 0.2296 |
| | 9 | 250.252044 | 0.00664 | 250.221130 | 0.0243 |
| | 10 | 250.290513 | 0.04511 | 250.363882 | 0.1185 |
| 512 | 8 | 250.206884 | 0.03852 | 250.058060 | 0.1873 |
| | 9 | 250.244890 | 0.00051 | 250.183410 | 0.0620 |
| | 10 | 250.283279 | 0.03788 | 250.356760 | 0.1114 |
| 1024 | 8 | 250.238977 | 0.00642 | 250.143770 | 0.1016 |
| | 9 | 250.244751 | 0.00065 | 250.268955 | 0.0236 |
| | 10 | 250.270855 | 0.02546 | 250.396782 | 0.1514 |
| 2048 | 8 | 250.237612 | 0.00779 | 250.103221 | 0.1422 |
| | 9 | 250.247710 | 0.00231 | 250.320873 | 0.0755 |
| | 10 | 250.275206 | 0.02981 | 250.347819 | 0.1024 |

*Маркираните редове представляват оптималните k стойности за всяко N .

В Таблица 1.7 са представени резултатите за оценяване на λ_{max} чрез *почти оптималните* и *класическите* СКМК алгоритми, базирани на разбъркани редица на Холтън. При всички алгоритми абсолютната грешка намалява с увеличаване на броя преходи k , като оптималният баланс се постига при $k = 12$. В този случай *почти оптималният* алгоритъм $PQMC_{(H)}$ достига абсолютна грешка от порядъка на 10^{-3} (например, 4.59×10^{-3} при $N = 2048$), докато *класическият* вариант $PQMC_{(H)}^{(n)}$ демонстрира значително по-големи грешки, достигащи 0.0726 при същия брой реализации. Това ясно показва предимството на оптимизираната конструкция на алгоритъма. В Таблица 1.8 са представени аналогични резултати за алгоритмите, използващи редици на Соболев. *Почти оптималният* алгоритъм $PQMC_{(S)}$ демонстрира най-висока точност, като в балансирания случай при $k = 12$ абсолютната грешка е от порядъка на 10^{-4} (например, 5.4×10^{-4} при $N = 2048$). За разлика от него, *класическият* вариант $PQMC_{(S)}^{(n)}$ достига значително по-големи отклонения, като при $N = 256$ грешката е 0.3172. Това потвърждава по-добрата равномерност на редиците на Соболев и тяхното предимство при по-високи размерности.



Фигура 1.4: Графично сравняване на λ_{max} за матрицата A ($n = 500$) с приближените резултати за $\lambda_{max}^{(k)}$, получени при балансирания случай на грешките, т.е. когато ($k = 9, N \geq 128$). Приложени алгоритми: $PMC_{(MT)}$, $PMC_{(MS)}$, $PMC_{(MT)}^{(n)}$ и $PMC_{(MS)}^{(n)}$.

Таблица 1.7: Резултати от $PQMC_{(H)}$ и $PQMC_{(H)}^{(n)}$ при различни N и k за матрица A с размерност $n = 500$. Абсолютната грешка е за случая $\lambda_2/\lambda_1 \approx 0.0363$.

| N | k | $\lambda_1^{(k)}$ using | Absolute Error | $\lambda_1^{(k)}$ using | Absolute Error |
|------|-----------|-------------------------|----------------|-------------------------|--------------------|
| | | $PQMC_{(H)}$ | $PQMC_{(H)}$ | $PQMC_{(H)}^{(n)}$ | $PQMC_{(H)}^{(n)}$ |
| 128 | 11 | 250.167707 | 0.07769 | 251.486594 | 1.2412 |
| | 12 | 250.249391 | 0.00399 | 249.309133 | 0.9363 |
| | 13 | 250.357823 | 0.11242 | 251.999063 | 1.7537 |
| 256 | 11 | 250.197841 | 0.04756 | 249.470953 | 0.7744 |
| | 12 | 250.239655 | 0.00574 | 250.770441 | 0.5250 |
| | 13 | 250.357573 | 0.11217 | 251.711421 | 1.4660 |
| 512 | 11 | 250.089218 | 0.15618 | 249.591525 | 0.6539 |
| | 12 | 250.236033 | 0.00937 | 250.669970 | 0.4246 |
| | 13 | 250.369509 | 0.12411 | 250.978674 | 0.7333 |
| 1024 | 11 | 250.127756 | 0.11764 | 249.672942 | 0.5725 |
| | 12 | 250.250718 | 0.00532 | 250.090116 | 0.1553 |
| | 13 | 250.387049 | 0.14165 | 250.829267 | 0.5839 |
| 2048 | 11 | 250.166549 | 0.07885 | 249.987224 | 0.2582 |
| | 12 | 250.249986 | 0.00459 | 250.172833 | 0.0726 |
| | 13 | 250.321274 | 0.07587 | 250.581459 | 0.3361 |

*Маркираните редове представляват оптималните k стойности за всяко N .

Фигура 1.5 илюстрира сходимостта на *почти оптималните* и *класическите* СКМК алгоритми при оптималната стойност $k = 12$. *Почти оптималните* алгоритми $PQMC_{(S)}$ и $PQMC_{(H)}$ следват плътно точната стойност на λ_{max} за всички стойности на N , демонстрирайки по-бърза и по-гладка сходимост. За разлика от тях, *класическите* варианти се характеризират с по-бавна сходимост и по-големи колебания, особено при по-малки стойности на N .

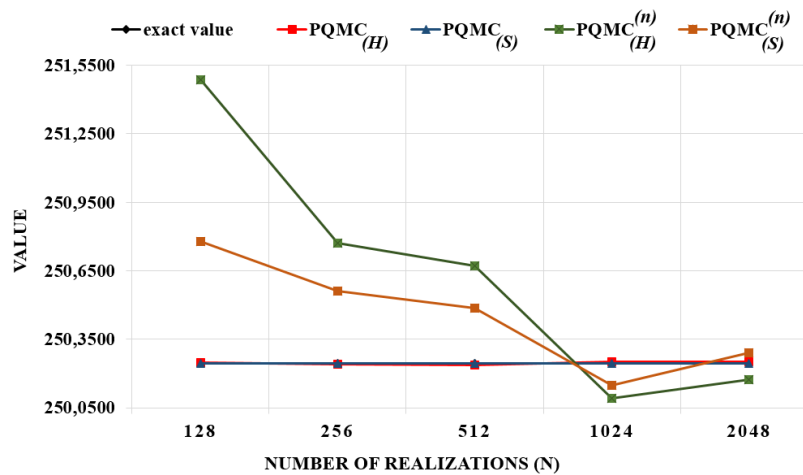
На Фигура 1.6 са показани сравненията между *почти оптималните* СКМК и СКМК алгоритми. *Почти оптималните* СКМК алгоритми $PMC_{(MT)}$ и $PMC_{(MS)}$ достигат минимални грешки съответно 3.8×10^{-4} и 6.5×10^{-4} при $N = 1024$ и $k = 9$, докато *почти оптималният* СКМК алгоритъм $PQMC_{(S)}$ осигурява съпоставима или по-висока точност при $k = 12$.

В Таблица 1.9 са сравнени емпиричните оценки за дисперсиите на случайната величина $\theta^{(k)}$ при *класическите* и *почти оптималните* СКМК алгоритми. Оценките са извършени за матрици с размерности $n = 100$ и $n = 500$. Резултатите показват, че при *почти оптималните* СКМК алгоритми дисперсията е с около два порядъка по-малка в сравнение с *класическите* им варианти. Тази оценка не зависи съществено от използвания генератор или от броя реализации N , което е в съзвучие с теоретичните резултати. Фактът, че дисперсията при *почти оптималните* версии е с два порядъка по-малка от *класическите*

Таблица 1.8: Резултати от $PQMC_{(S)}$ и $PQMC_{(S)}^{(n)}$ при различни N и k за матрица A с размерност $n = 500$. Абсолютната грешка е за случая $\lambda_2/\lambda_1 \approx 0.0363$.

| N | k | $\lambda_1^{(k)}$ using PQMC _(S) | Absolute Error PQMC _(S) | $\lambda_1^{(k)}$ using PQMC _(S) ⁽ⁿ⁾ | Absolute Error PQMC _(S) ⁽ⁿ⁾ |
|------|-----------|--|---------------------------------------|---|--|
| 128 | 11 | 250.236415 | 0.00898 | 249.415659 | 0.8297 |
| | 12 | 250.245037 | 0.00036 | 250.777310 | 0.5319 |
| | 13 | 250.330000 | 0.08460 | 250.989461 | 0.7441 |
| 256 | 11 | 250.199189 | 0.04621 | 249.483392 | 0.7620 |
| | 12 | 250.245119 | 0.00028 | 250.562622 | 0.3172 |
| | 13 | 250.291367 | 0.04597 | 250.704555 | 0.4592 |
| 512 | 11 | 250.099484 | 0.14592 | 249.632500 | 0.6129 |
| | 12 | 250.244872 | 0.00053 | 250.486175 | 0.2408 |
| | 13 | 250.378047 | 0.13265 | 250.679597 | 0.4342 |
| 1024 | 11 | 250.140000 | 0.10540 | 249.811533 | 0.4339 |
| | 12 | 250.244996 | 0.00040 | 250.148126 | 0.0973 |
| | 13 | 250.351139 | 0.10574 | 250.651883 | 0.4065 |
| 2048 | 11 | 250.193656 | 0.05174 | 249.937040 | 0.3084 |
| | 12 | 250.245937 | 0.00054 | 250.290864 | 0.0455 |
| | 13 | 250.272562 | 0.02716 | 250.555662 | 0.3103 |

*Маркираните редове представляват оптималните k стойности за всяко N .

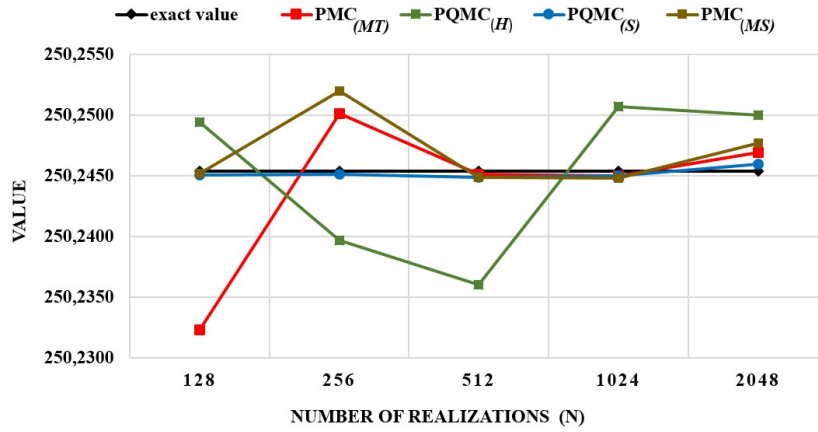


Фигура 1.5: Графично сравняване на λ_{max} за матрицата A ($n = 500$) с приближените резултати за $\lambda_{max}^{(k)}$, получени при балансиран случай на грешките, т.е. когато ($k = 12, N \geq 128$). Приложени алгоритми: $PQMC_{(S)}$, $PQMC_{(H)}$, $PQMC_{(S)}^{(n)}$ и $PQMC_{(H)}^{(n)}$.

Таблица 1.9: Дисперсия на оценките на максималната собствена стойност при почти оптималния и класическия СМК алгоритъм за различни размерности на матрицата, брой реализации (N), брой преходи (k) и използваните генератори на псевдослучайни числа (MT и MS)

| matrix size | N | k | $Var(PMC_{(MT)})$ | $Var(PMC_{(MT)}^{(n)})$ | $Var(PMC_{(MS)})$ | $Var(PMC_{(MS)}^{(n)})$ |
|-------------|------|---|-------------------|-------------------------|-------------------|-------------------------|
| 100 | 512 | 8 | 1.169203e-06 | 1.720495e-04 | 1.236100e-06 | 1.781005e-04 |
| | 2048 | 8 | 1.300635e-06 | 1.772653e-04 | 1.326996e-06 | 1.792424e-04 |
| 500 | 512 | 9 | 1.186151e-08 | 1.229778e-05 | 1.269925e-08 | 1.285468e-05 |
| | 2048 | 9 | 1.189763e-08 | 1.184918e-05 | 1.268721e-08 | 1.299595e-05 |

версии, потвърждава отново направените изводи, че тези алгоритми осигуряват по-точно оценяване на максималната собствена стойност на симетрични матрици при намалена изчислителна сложност.



Фигура 1.6: Графично сравнение на λ_{max} на симетричната матрица A ($n = 500$) с $\lambda_{max}^{(k)}$ в балансиращите случаи. Използвани са *почти оптималните* СМК и СКМК алгоритми.

1.5 Обобщение

В тази глава е изследвана ефективността на *почти оптималните* СМК и СКМК алгоритми за оценка на най-голямата собствена стойност на симетрични плътни матрици за $n = 100$ и $n = 500$. Числените експерименти показват, че оптимизираните алгоритми значително превъзхождат *класическите* по точност и изчислителна сложност чрез целенасочен избор на преходната матрица и балансиране на систематичната и стохастичната грешка. При СМК алгоритмите оптималният брой преходи е $k = 8$ за $n = 100$ и $k = 9$ за $n = 500$, като генераторът Mersenne Twister (MT) осигурява по-малка дисперсия от Middle Square (MS). При СКМК оптималният брой преходи е $k = 11$ за $n = 100$ и $k = 12$ за $n = 500$, като разбърканите редици на Соболев водят до по-гладка сходимост и по-малки грешки от тези на Холтън. Абсолютните грешки при *почти оптималните* алгоритми са от порядъка на 10^{-4} , докато при *класическите* достигат до 10^{-2} , потвърждавайки значителното намаляване на дисперсията и повишаване на точността.

В заключение, *почти оптималните* СМК и СКМК алгоритми осигуряват по-висока точност при същата изчислителна сложност, което ги прави предпочитан избор за практически приложения. При тяхното конструиране трябва да се съобразяваме и с избора на генераторите на псевдослучайни числа и от избора на редиците с малък дискрепанс. Описаните резултати са публикувани в [36, 37].

2. Стохастичен Резолвентен метод: разработване и изследване на ефективни алгоритми за оценяване на минимална собствена стойност

2.1 Постановка на задачата

Нека $A = \{a_{ij}\}_{i,j=1}^n \in \mathbb{R}^{n \times n}$ е несингулярна симетрична матрица. Собствените стойности на A се определят като всички реални стойности на параметъра $\lambda(A)$, за които уравнението:

$$A\mathbf{x} = \lambda(A)\mathbf{x}, \quad \text{където } \mathbf{x} \in \mathbb{R}^n, \quad (2.1)$$

има ненулево решение. Приема се, че собствените стойности удовлетворяват наредбата:

$$0 < \lambda_{min} = \lambda_n < \lambda_{n-1} \leq \lambda_{n-2} \leq \dots \leq \lambda_2 < \lambda_1 = \lambda_{max}. \quad (2.2)$$

като минималната собствена стойност λ_{min} е положителна и съгласно (2.2) е единствена.

Разглеждаме задачата за разработване и численото изследване на ефективни стохастични алгоритми за нейното оценяване.

2.2 Методи за изчисляване на минималната собствена стойност

Нека $R = \{r_{ij}\}_{i,j=1}^n \in \mathbb{R}^{n \times n}$ е несингулярна симетрична матрица, чиито собствени стойности удовлетворяват наредбата:

$$\mu_{min} = |\mu_n| < |\mu_{n-1}| \leq |\mu_{n-2}| \leq \dots \leq |\mu_2| < |\mu_1| = \mu_{max}, \quad (2.3)$$

и за тях уравнението:

$$R\mathbf{y} = \mu(R)\mathbf{y} \quad (2.4)$$

има ненулево решение. От теорията на линейната алгебра е известно, че при определени зависимости между матриците R и A съществуват връзки между техните собствени стойности и вектори. Разглеждат се три основни случая:

- При обратна матрица $R = A^{-1}$ собствените стойности са свързани чрез $\mu = \frac{1}{\lambda}$. Това позволява използване на *Обратния Степенен* метод (англ. *Inverse Power method*) за намиране на минималната собствена стойност на A , но изисква скъпо изчисляване на обратна матрица.
- При отместена матрица $R_q = A - q\mathbf{I}$ е валидно $\mu = \lambda - q$. Това води до *Отместен Степенен* метод (англ. *Shifted Power method*), при който изборът на параметъра q влияе върху оценката на минималната собствена стойност.
- При резолвентна матрица $R_q = (\mathbf{I} - qA)^{-1}$ собствените стойности са свързани чрез $\mu = \frac{1}{1 - q\lambda}$, а съответният метод се нарича *Резолвентен Степенен* метод (англ. *Resolvent Power method*). Матрицата R_q може да се представи като граница от безкраен сходящ ред при условие $|q\lambda| < 1$ или еквивалентно $|q|||A|| < 1$, а именно:

$$R_q = \sum_{i=0}^{\infty} q^i A^i = [\mathbf{I} - qA]^{-1}. \quad (2.5)$$

В зависимост от знака на параметъра q се получават различни връзки за максималната собствена стойност:

$$\mu_{max} = \frac{1}{1 - q\lambda_{min}}, \quad \text{при } q < 0, \quad (2.6)$$

$$\mu_{max} = \frac{1}{1 - q\lambda_{max}}, \quad \text{при } q > 0. \quad (2.7)$$

В рамките на настоящото изследване основен интерес представлява случаят $q < 0$, който позволява ускорено приближение и ефективно оценяване на екстремните собствени стойности.

2.2.1 Резолвентен Степенен метод

В *класическия* Степенен метод най-голямата собствена стойност λ_{max} на симетричната матрица A се представя като граница на коефициента на Рейли (1.4), което позволява практическо приближение чрез крайни степени. На тази основа се въвежда Резолвентният Степенен метод, при който матрицата A се заменя с резолвентната матрица $R_q = (\mathbf{I} - qA)^{-1}$, която съществува при условие $|q|||A|| < 1$. Резолвентната матрица има развитие в Нойманов ред, а нейната степен R_q^m може да се представи чрез обобщен биномен ред:

$$R_q^m = (\mathbf{I} - qA)^{-m} = \sum_{i=0}^{\infty} q^i C_{i+m-1}^i A^i, \quad C_{i+m-1}^i = \frac{(m+i-1)!}{(m-1)!i!}, \quad i = 0, 1, 2, \dots \quad (2.8)$$

Безкрайният ред за R_q^m (q се нарича параметър за ускорение) може да се запише като:

$$R_q^m = (I - qA)^{-m} = \mathbf{I} + mqA + \frac{m(m+1)}{2}q^2A^2 + \frac{m(m+1)(m+2)}{6}q^3A^3 + \dots \quad (2.9)$$

Прилагането на Степенния метод към резолвентната матрица (R_q) води до Резолвентния Степенен метод:

$$\lim_{m \rightarrow \infty} \frac{(\mathbf{h}, R_q^m \mathbf{f})}{(\mathbf{h}, R_q^{m-1} \mathbf{f})} = \lim_{m \rightarrow \infty} \frac{(\mathbf{h}, (\mathbf{I} - qA)^{-m} \mathbf{f})}{(\mathbf{h}, (\mathbf{I} - qA)^{-(m-1)} \mathbf{f})} = \mu_{max} = \frac{1}{1 - q\lambda}, \quad (2.10)$$

който позволява намирането на максималната собствена стойност μ_{max} на резолвентната матрица и съответно оценка на екстремалните собствени стойности на A . В зависимост от знака на параметъра q , методът може да се използва за приближено пресмятане както на λ_{max} , така и на λ_{min} , съгласно формулите [19, 20, 22]:

$$\lambda = \frac{(\mathbf{h}, AR_q^m \mathbf{f})}{(\mathbf{h}, R_q^m \mathbf{f})} \approx \lambda_{max}, \quad \text{ако } q > 0, \quad (2.11)$$

$$\frac{(\mathbf{h}, AR_q^m \mathbf{f})}{(\mathbf{h}, R_q^m \mathbf{f})} \approx \frac{1}{q} \left(1 - \frac{1}{\mu^{(m)}}\right) \approx \lambda_{min}, \quad \text{ако } q < 0. \quad (2.12)$$

В практиката резолвентната матрица се апроксимира чрез прекъсване на реда до краен брой членове k :

$$R_{q,k}^m = \sum_{i=0}^k C_{i+m-1}^i q^i A^i, \quad (2.13)$$

което води до систематична грешка, равна на "опашката" на реда:

$$\Delta_k^{(m)}(q) = (I - qA)^{-m} - R_{q,k}^m = \sum_{i=k+1}^{\infty} C_{i+m-1}^i q^i A^i. \quad (2.14)$$

Получени са оценки за тази грешка и условия за избор на параметрите m и k , така че тя да бъде по-малка от зададено ε , т.е.:

$$|\Delta_k^{(m)}(q)| < \varepsilon. \quad (2.15)$$

В сила е следната оценката:

$$|\Delta_k^{(m)}(t)| \leq T_{m,k}(t) = \sum_{i=k+1}^{\infty} C_{i+m-1}^i t^i, \quad t = |q||\lambda_{max}| < 1. \quad (2.16)$$

Съществува линейна връзка между параметрите m и k при условие $t = |q||\lambda_{max}| < 1$. Следните оценки са получени за връзката между k и m :

- Неравенство за точност:

$$T_{m,k}(t) \leq \frac{b_{k+1}}{1 - \alpha} = \frac{\binom{m+k}{k+1} t^{k+1}}{1 - \frac{m+k+1}{k+2} t} < \varepsilon, \quad b_i = C_{i+m-1}^i t^i, \quad \alpha := \frac{m+k+1}{k+2}, \quad t < 1. \quad (2.17)$$

- Линейни зависимости между параметрите:

$$m < \frac{1-t}{t}(k+1) + \frac{1}{t} \quad \text{или} \quad m_{max} = \left\lfloor \frac{(1-t)k+2}{t} - 1 \right\rfloor. \quad (2.18)$$

- Асимптотична връзка (ако m расте, а t е фиксирано):

$$k \gtrsim \frac{t}{1-t}m, \quad \text{за големи стойности на } m. \quad (2.19)$$

- Минимална стойност на k при фиксирано t се определя от формулата:

$$k_{min} = \left\lceil \frac{tm-1}{1-t} - 1 \right\rceil. \quad (2.20)$$

Получените зависимости позволяват ефективен избор на параметри при числена реализация и са основа за построяване на Стохастичен Резолвентен метод, когато се търси баланс между систематична и стохастична грешка.

2.2.2 Стохастичен Резолвентен метод

Стохастичният Резолвентен метод, въведен от Караиванова и Димов [20, 25], представлява обобщение на Резолвентния Степенен метод за оценяване на екстремалните собствени стойности на матрица A . Основен елемент е конструкцията на случайните величини $\theta^{(i)}$, дефинирани чрез верига на Марков, чиито математически очаквания съвпадат със скаларните произведения $(\mathbf{h}, A^i \mathbf{f})$, $i = 1, 2, 3, \dots$, т.е.:

$$E[\theta^{(i)}] = (\mathbf{h}, A^i \mathbf{f}), \quad i = 1, 2, 3, \dots \quad (2.21)$$

При условие за параметъра q , удовлетворяващ $|q \lambda_{max}| < 1$, е доказано в [17], че скаларното произведение с резолвентната матрица може да се представи като безкраен ред:

$$(\mathbf{h}, R_q^m \mathbf{f}) = (\mathbf{h}, (\mathbf{I} - qA)^{-m} \mathbf{f}) = \sum_{i=0}^{\infty} q^i C_{i+m-1}^i E[\theta^{(i)}], \quad (2.22)$$

което позволява стохастично оценяване чрез Монте Карло методи. На практика математическите очаквания се апроксимират чрез средната стойност:

$$\bar{\theta}_N^{(i)} = \frac{1}{N} \sum_{s=1}^N (\theta^{(i)})_s, \quad i = 1, 2, 3, \dots \quad (2.23)$$

като стохастичната грешка е от порядък $\mathcal{O}(N^{-1/2})$. При фиксирана точност ε е решена задачата за определяне на минималния брой реализации N , необходим за гарантиране на грешка под зададен праг и при условие за ограничена дисперсия и $\|A\| < 1$:

$$N \geq \left\{ \frac{C_k \sigma}{\varepsilon} \right\}^2. \quad (2.24)$$

Поради прекъсване на безкрайния ред до краен брой членове k възникват две грешки: (1) стохастична (Монте Карло апроксимация) и (2) систематична (от прекъсването на реда). Изборът на параметрите m, k контролира систематичната грешка, а N - стохастичната. Балансът между тях е решаващ за точността и изчислителната ефективност. Приближената формула за собствена стойност е:

$$\lambda \approx \frac{E[\theta^{(1)}] + qC_m^1 E[\theta^{(2)}] + q^2 C_{m+1}^2 E[\theta^{(3)}] + \dots + q^k C_{k+m-1}^k E[\theta^{(k+1)}]}{1 + qC_m^1 E[\theta^{(1)}] + q^2 C_{m+1}^2 E[\theta^{(2)}] + \dots + q^k C_{k+m-1}^k E[\theta^{(k)}]}, \quad (2.25)$$

като знакът на параметъра q определя дали се оценява максималната (λ_{max} при $q > 0$) или минималната собствена стойност (λ_{min} при $q < 0$). При използване на псевдослучайни числа се получава Резолвентен Монте Карло (РМК) метод, а при редици с малък дискрепанс - Резолвентен Квази-Монте Карло (РКМК). Конструкцията на случайните величини се базира на подходящ избор на начална плътност и матрица на преходните вероятности, като числените резултати показват по-висока точност спрямо *класическите* детерминистични алгоритми.

2.3 Почти оптимален Резолвентен Монте Карло и Резолвентен Квази-Монте Карло алгоритми

Нека матрицата A се представя чрез редовете \mathbf{a}_i , като се въвеждат L_1 -норми за вектори и матрици, които служат като основа при конструиране на *почти оптимални* стохастични резолвентни алгоритми. Изборът на начален вектор на плътността и матрица на преходните вероятности се реализира чрез нормализиране на абсолютните стойности на компонентите на \mathbf{h} и редовете на матрицата A , което води до *почти оптимална* структура на веригата на Марков.

Представен е *почти оптимален* Резолвентен Квази-Монте Карло (РКМК) алгоритъм за оценяване на минималната собствена стойност на симетрична матрица. Алгоритъмът използва редици с малък дискрепанс (например, разбъркана редица на Соболю) за генериране на реализации във веригата на Марков и изчисляване на случайните величини $\theta^{(i)}$. При използване на псевдослучайни генератори методът се нарича *почти оптимален* Резолвентен Монте Карло (РМК) алгоритъм. В стъпка 4 на алгоритъма се прилага разбъркана редица на Соболю за изчисляване стойностите на случайните величини $\theta^{(i)}$, $i = 1, 2, \dots, k + 1$, чрез верига на Марков.

Алгоритъм 2, представен чрез псевдокод, описва *почти оптималния* Резолвентен Квази-Монте Карло (РКМК) алгоритъм (англ. *almost optimal Resolvent Quasi-Monte Carlo (RQMC) algorithm*) за оценяване на минималната собствена стойност на симетричната матрица A . В случай, че в стъпка 4 се използват генератори на псевдослучайни числа (PRNGs), алгоритъмът се определя като *почти оптимален* Резолвентен Монте Карло (РМК) алгоритъм (англ. *almost optimal Resolvent Monte Carlo (RMC) algorithm*). Псевдокодът, представящ реализацията на *почти оптималния* РМК алгоритъм, ще означаваме с **Алгоритъм 2а**. Двата алгоритъма принадлежат към класа на итерационните стохастични методи и се характеризират с наличието на два вида грешки: стохастична и систематична. Стохастичната грешка при МК методите е от порядък $\mathcal{O}(N^{-1/2})$, докато при КМК методите (базирани на редици с малък дискрепанс) - $\mathcal{O}(N^{-1}(\log N)^k)$. При използване на редици с малък дискрепанс в пространства с висока размерност може да доведе до корелации между координатите на квазислучайните редици, което се преодолява чрез рандомизирането (разбъркването) им.

Algorithm 2 Pseudocode for computing the almost optimal Resolvent QMC algorithm

- 1: INPUT: matrix $A = \{a_{ij}\} \in \mathbb{R}^{n \times n}$; vectors $\mathbf{h}, \mathbf{f} \in \mathbb{R}^n$, positive integers n, N, k, m and the real number q , where $(-1)/\|A\| < q < 0$
- 2: COMPUTE: $p_i = \frac{|h_i|}{\sum_{i=1}^n |h_i|}$; $p_{ij} = \frac{|a_{ij}|}{\sum_{j=1}^n |a_{ij}|}$, $1 \leq i \leq n$, $1 \leq j \leq n$
- 3: COMPUTE: $a_i = \sum_{j=1}^n |a_{ij}|$, $1 \leq i \leq n$, $\|\mathbf{h}\| = \sum_{i=1}^n |h_i|$
- 4: GENERATE: N elements of the $(k + 1)$ - dimensional Sobol sequence
- 5: CONSTRUCT: N realizations of the Markov chain with integer elements

$$l_1^{(s)} \rightarrow \dots \rightarrow l_i^{(s)} \rightarrow \dots \rightarrow l_k^{(s)} \rightarrow l_{k+1}^{(s)}, \quad 1 \leq l_i^{(s)} \leq n, \quad 1 \leq i \leq k + 1, \quad 1 \leq s \leq N$$

- 6: COMPUTE:

$$\bar{\theta}^{(i)} = \frac{1}{N} \|\mathbf{h}\| \sum_{s=1}^N \text{sign}\{h_{l_1^{(s)}} \prod_{i=1}^{k+1} (a_{l_i^{(s)} l_{i+1}^{(s)}})\} \prod_{i=1}^{k+1} a_{l_i^{(s)}} f_{l_{i+1}^{(s)}}; \quad 1 \leq i \leq k + 1$$

- 7: COMPUTE:

$$C_{i+m-1}^i = \frac{(i+m-1)!}{i!(m-1)!}, \quad \text{where } 1 \leq i \leq k$$

- 8: OUTPUT:

$$\lambda_{\min} \approx \frac{\bar{\theta}^{(1)} + qC_m^1 \bar{\theta}^{(2)} + q^2 C_{m+1}^2 \bar{\theta}^{(3)} + \dots + q^k C_{k+m-1}^k \bar{\theta}^{(k+1)}}{1 + qC_m^1 \bar{\theta}^{(1)} + q^2 C_{m+1}^2 \bar{\theta}^{(2)} + \dots + q^k C_{k+m-1}^k \bar{\theta}^{(k)}}$$

Систематичната грешка произтича от използването на крайна (прекъсната) верига на Марков и е аналогична на прекъсната верига на Нойман при детерминистичните методи. Тя зависи от параметъра на прекъсване k в представянето на резолвентната матрица $R_{q,k}^m$ като граница на безкраен ред при условие $|q|\|A\| < 1$. Параметърът m определя степента на резолвентната матрица и влияе върху сходимостта на метода:

$$\mathcal{O}\left(\left|\frac{\mu_2}{\mu_1}\right|^m\right), \quad (2.26)$$

където μ_1, μ_2 са най-големите собствени стойности на резолвентната матрица. Изразено чрез собствените стойности на A :

$$\mathcal{O}\left(\left|\frac{1-q\lambda_1}{1-q\lambda_2}\right|^m\right), \text{ при } q > 1 \text{ и } \mathcal{O}\left(\left|\frac{1-q\lambda_n}{1-q\lambda_{n-1}}\right|^m\right), \text{ при } q < 1. \quad (2.27)$$

Изборът на ускоряващия параметър q определя дали алгоритъмът оценява минималната ($q \in (-\frac{1}{\|A\|}, 0)$) или максималната собствена стойност ($q > 0$). Изчислителният процес включва генериране на N реализации на веригата на Марков, пресмятане на средните стойности при симулиране на случайните величини и използване на биномни коефициенти за оценка на собствената стойност.

Балансът между стохастичната и систематичната грешка е съществен за постигане на предварително зададена точност ε при оптимална изчислителна цена. Този баланс се контролира чрез параметрите N (брой реализации), k (дължина на веригата на Марков), m (степената на резолвентната матрица) и q (ускоряващ параметър). Числените експерименти показват, че е необходимо изборът на тези параметри да бъде съгласуван с получените теоретични оценки за сходимост.

2.4 Числени резултати и дискусия

Числените резултати са постигнати чрез прилагането на *почти оптималния* Резолвентен Квази-Монте Карло (РКМК) алгоритъм (**Алгоритъм 2**) и *почти оптималния* Резолвентен Монте Карло (РМК) алгоритъм (**Алгоритъм 2а**) за оценяване на минималната собствена стойност на случайно генерирани симетрични матрици с размерности $n = 500$ и $n = 1000$. *Почти оптималният* РКМК алгоритъм е реализиран с генератора BRODA's Sobol RSG, който позволява генериране на разбъркани редици на Собол с необходимата размерност при конструиране на веригата на Марков. *Почти оптималният* РМК алгоритъм (**Алгоритъм 2а**) е реализиран с генератора на псевдослучайни числа Mersenne Twister (MT). Двата алгоритъма са имплементирани на C++, като предварително се използва подпрограма за генериране на случайна симетрична матрица A със спектър в единичната окръжност, удовлетворяващ условието:

$$0 < c < \lambda_n < \lambda_{n-1} \leq \dots \leq \lambda_2 < \lambda_1 < 1,$$

където константата c е избрана да бъде $c = 0.15$. Това гарантира положителност и спектрална отделимост, необходими за сходимостта на Резолвентния Степенен метод.

Разгледани са два тестови случая:

- (i) Матрица A с размерност $n = 500$, като двете най-големи и двете най-малки собствени стойности са: $\lambda_1 = 0.620713$, $\lambda_2 = 0.335683$, $\lambda_{499} = 0.15452$ и $\lambda_{500} = 0.152992$.
- (ii) Матрица A с размерност $n = 1000$, като двете най-големи и двете най-малки собствени стойности са: $\lambda_1 = 0.688181$, $\lambda_2 = 0.443078$, $\lambda_{999} = 0.154138$ и $\lambda_{1000} = 0.151911$.

Различието между двете най-големи и двете най-малки собствени стойности осигурява надеждна сходимост на метода при оценяване на λ_{min} .

Числените експерименти са проведени върху високопроизводителен клъстер (ИИКТ-БАН), състоящ се от 12 броя сървъри Fujitsu Primergy RX 2540 M4 като всеки от тях е със следната конфигурация: 128 GB оперативна памет; два броя процесори - CPU 2x Intel Xeon Gold 5118 2.30 GHz с 24 ядра; 2 броя по 800 GB SSD; 36 TB HDD, Графична карта - NVIDIA Tesla V100 с 32 GB оперативна памет. Сървърите са свързани с InfiniBand за бърз обмен на данни с малка латентност. Операционната система е Red Hat Enterprise Linux 7.6 и NVIDIA GRID.

Реализацията на **Алгоритъм 2** е хибридна: генерирането на разбърканите редици на Собол се извършва върху GPU (NVIDIA Tesla V100, 5120 CUDA ядра), докато останалите изчисления се изпълняват върху CPU. **Алгоритъм 2а** е реализиран върху CPU с генератора Mersenne Twister.

Произволните вектори \mathbf{h} и \mathbf{f} са избрани с L_1 -норма по следния начин:

$$\|\mathbf{h}\| = \|\mathbf{f}\| = \sum_{i=1}^n \frac{1}{n} = 1,$$

за да се осигури коректна нормировка при стохастичните оценки. Числените експерименти показват, че при големи стойности на броя реализации N , времето за изпълнение на РКМК алгоритъма ($RQMC_{(S)}$) е

съществено по-малко от това на РМК алгоритъма ($RMC_{(MT)}$), което демонстрира предимството на Квази-Монте Карло подхода в комбинация с GPU ускорение при задачи с големи размерности за оценяване на минималната собствена стойност.

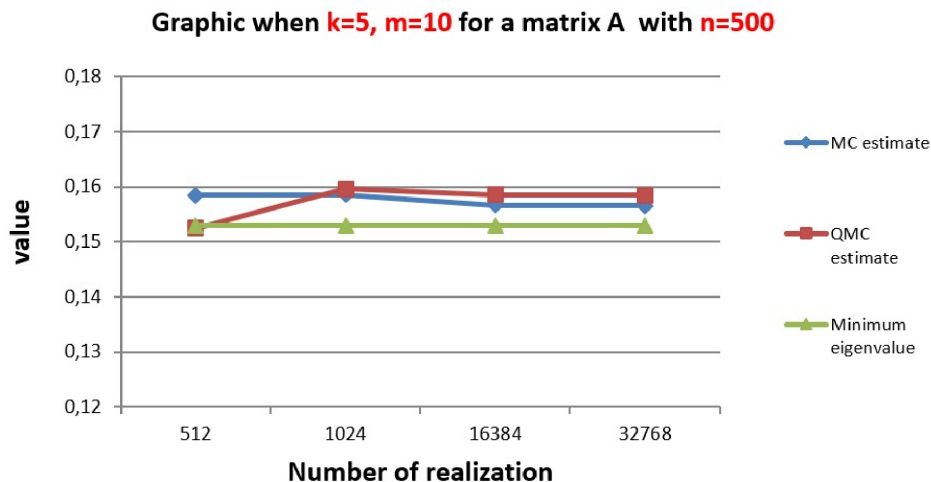
2.4.1 Случай (i): Размер на матрицата $n=500$

В Таблица 2.1 са представени числени резултати за оценяване на минималната собствена стойност на симетрична матрица с размерност $n = 500$, получени чрез *почти оптималните* алгоритми ($RMC_{(MT)}$ и $RQMC_{(S)}$) при различни стойности на параметрите N, k и m . В таблицата са включени и съответните абсолютни грешки на получените оценки. Параметърът на ускорение е фиксиран на $q = -0.157428$, а систематичната грешка е предварително зададена на ниво $\varepsilon_{sys} = 0.001$.

Таблица 2.1: Числени резултати за матрицата $A \in \mathbb{R}^{500 \times 500}$ при различни стойности на N, k и m с параметър на ускорение $q = -0.157428$. Точната минималната собствена стойност $\lambda_{\min} = \lambda_{500} = 0.152992$.

| N | k | m | λ_{\min} | Absolute error | λ_{\min} | Absolute error |
|-------------|-----|-----|------------------|------------------|------------------|------------------|
| | | | $RMC_{(MT)}$ | $RMC_{(MT)}$ | $RQMC_{(S)}$ | $RQMC_{(S)}$ |
| $512 * 2^0$ | 5 | 5 | 0.1585637 | 0.0055717 | 0.1561046 | 0.0031126 |
| | 5 | 10 | 0.1585163 | 0.0055243 | 0.1524884 | 0.0005036 |
| | 10 | 5 | 0.1556132 | 0.0026212 | 0.1560595 | 0.0030675 |
| $512 * 2^1$ | 5 | 5 | 0.158395 | 0.005403 | 0.1596503 | 0.0066583 |
| | 5 | 10 | 0.1585518 | 0.0055598 | 0.1595946 | 0.0066026 |
| | 10 | 5 | 0.1580794 | 0.0050874 | 0.1596616 | 0.0066696 |
| $512 * 2^5$ | 5 | 5 | 0.1572588 | 0.0042668 | 0.1591139 | 0.0061219 |
| | 5 | 10 | 0.1566524 | 0.0036604 | 0.1585388 | 0.0055468 |
| | 10 | 5 | 0.156875 | 0.003883 | 0.1591292 | 0.0061372 |
| $512 * 2^6$ | 5 | 5 | 0.157129 | 0.004137 | 0.1589384 | 0.0059464 |
| | 5 | 10 | 0.1565796 | 0.0035876 | 0.1584563 | 0.0054643 |
| | 10 | 5 | 0.1570216 | 0.0040296 | 0.158961 | 0.005969 |

Анализът на резултатите се основава на балансирането на двата основни източника на грешка - систематична и стохастична. Систематичната грешка произтича от прекъсването на Ноймановия ред при $i = k+1$, което съответства на краен брой преходи във веригата на Марков, както и от избора на степента m на резолвентната матрица. Параметрите m и k се избират въз основа на теоретични оценки за опаската на реда $T_{m,k}(t)$, при $t = |q||\lambda_1| < 1$, така че да е изпълнено условието $T_{m,k}(t) \leq \varepsilon_{sys} = 0.001$. По този начин систематичната грешка се контролира предварително и остава от порядък $\mathcal{O}(10^{-3})$.



Фигура 2.1: Сходимост към точната минимална собствена стойност на матрицата A при използване на РМК и РКМК при $k = 5, m = 10$ за различни на брой N

Стохастичната грешка възниква от апроксимацията на математическите очаквания чрез краен брой реализации N . При Монте Карло подхода ($RMC_{(MT)}$) стохастичната грешка е от порядък $\mathcal{O}(N^{-1/2})$ при

фиксирана доверителна вероятност. При Квази-Монте Карло подхода ($RQMC_{(S)}$) се очаква по-бърза сходимост. За оценяване на грешката се използват рандомизирани редици на Соболев. Стохастичната грешка зависи не само от N , но и от параметъра k , тъй като при нарастване на k се увеличава броят на оценяваните случайни величини $\theta^{(i)}$, $i = 1, \dots, k + 1$. Това води до натрупване на дисперсии както в числителя, така и в знаменателя на дробно-рационалната формула за оценяване на собствената стойност, което влияе върху необходимия брой реализации. Оптималният баланс изисква избор на такова N , че стохастичната грешка да бъде от същия порядък като предварително зададената систематична грешка ε_{sys} .

Получените резултати за $n = 500$ потвърждават теоретичните изводи. При подходящ избор на параметрите (m, k) , гарантиращ ограничаване на систематичната грешка под $\mathcal{O}(10^{-3})$, се постига висока точност още при сравнително малък брой реализации N , съпоставим с размерността на матрицата. Най-добра точност е получена при $k = 5, m = 10$ и $N = 512$, където абсолютната грешка е от същия порядък като предварително зададената систематична грешка. Увеличаването на N (например, до $N = 16384$ или $N = 32768$, виж Фигура 2.1 и Таблица 2.1) не води до монотонно намаляване на абсолютната грешка. Наблюдава се насищане или колебливо поведение, което е характерно за ситуации, в които систематичната грешка вече определя долната граница на точността. Допълнително влияние оказват числени ефекти, свързани с нарастването на биномните коефициенти при по-големи стойности на m и k , които увеличават чувствителността на дробно-рационалната формула към малки флуктуации в оценките на $\theta^{(i)}$. Следователно, увеличаването на N не гарантира автоматично повишаване на точността в случаите, когато систематичната компонента е определяща.

2.4.2 Случай (ii): Размер на матрицата $n=1000$

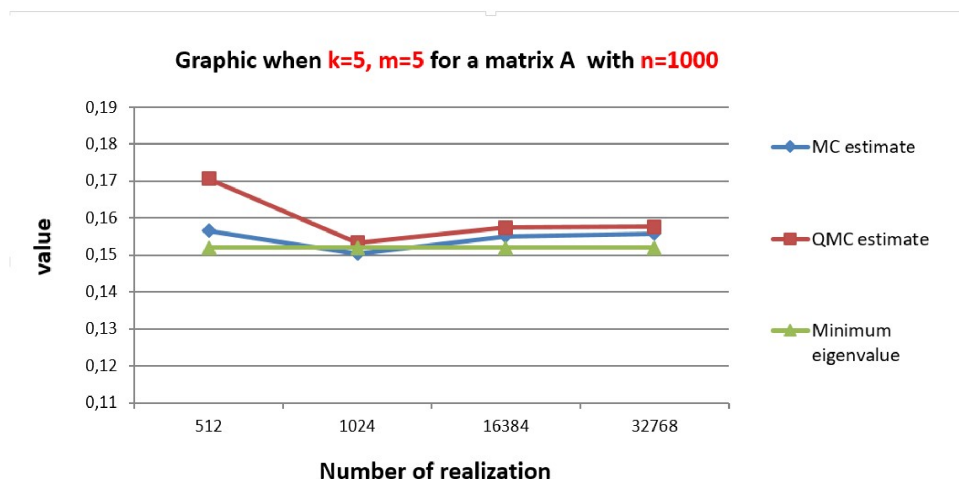
В Таблица 2.2 са представени числени резултати за оценяване на минималната собствена стойност на симетрична матрица с размерност $n = 1000$, получени чрез *почти оптималните* алгоритми $RM C_{(MT)}$ и $RQMC_{(S)}$, заедно със съответните абсолютни грешки. В тези експерименти параметърът на ускорение е фиксиран на $q = -0.2302458$, като абсолютната му стойност е по-голяма в сравнение със случая $n = 500$. Това води до увеличаване на параметъра $t = |q||\lambda_1|$ и налага, при същите стойности на k и m , систематичната грешка да бъде фиксирана на по-високо ниво $\varepsilon_{sys} = 0.01$, за да бъдат изпълнени теоретичните условия за оценка на опашката на реда.

Таблица 2.2: Числени резултати за матрицата $A \in \mathbb{R}^{1000 \times 1000}$ при различни стойности на N, k и m с параметър на ускорение $q = -0.2302458$. Точната минимална стойност $\lambda_{\min} = \lambda_{1000} = 0.151911$.

| N | k | m | λ_{\min} | Absolute error | λ_{\min} | Absolute error |
|-------------|-----|-----|------------------|------------------|------------------|------------------|
| | | | $RM C_{(MT)}$ | $RM C_{(MT)}$ | $RQMC_{(S)}$ | $RQMC_{(S)}$ |
| $512 * 2^0$ | 5 | 5 | 0.1556059 | 0.0045949 | 0.1705751 | 0.0186641 |
| | 5 | 10 | 0.168251 | 0.01634 | 0.1886959 | 0.0367849 |
| | 10 | 5 | 0.1593428 | 0.0074318 | 0.1701307 | 0.0182197 |
| $512 * 2^1$ | 5 | 5 | 0.1504153 | 0.0014957 | 0.1533797 | 0.0014687 |
| | 5 | 10 | 0.1417805 | 0.0101305 | 0.1499974 | 0.0019136 |
| | 10 | 5 | 0.156739 | 0.004828 | 0.1534383 | 0.0015273 |
| $512 * 2^5$ | 5 | 5 | 0.1550015 | 0.0030905 | 0.1573842 | 0.0054732 |
| | 5 | 10 | 0.1495483 | 0.0023627 | 0.1539357 | 0.0020247 |
| | 10 | 5 | 0.1559548 | 0.0040438 | 0.1575454 | 0.0056344 |
| $512 * 2^6$ | 5 | 5 | 0.1557509 | 0.0038399 | 0.1576339 | 0.0057229 |
| | 5 | 10 | 0.1537008 | 0.0017898 | 0.1562911 | 0.0043801 |
| | 10 | 5 | 0.1562905 | 0.0043795 | 0.15764225 | 0.0057312 |

Получените резултати (Таблица 2.2 и Фигура 2.2) показват, че най-добрият баланс между систематичната и стохастичната компонента на грешката се постига при $k = 5, m = 5$ и $N = 1024$. При тази комбинация абсолютната грешка е минимална и от същия порядък, като предварително се задава систематичната грешка. Това потвърждава теоретичния извод, че параметрите m и k , определени въз основа на фиксираното ниво на систематичната грешка, осигуряват необходимата точност на апроксимацията на резолвентната матрица. Броят на реализациите N трябва да бъде избран така, че стохастичната грешка да бъде съпоставима по порядък със систематичната. При по-големи стойности на m, k и N се наблюдава известно подобрене на абсолютната грешка по отношение на константния множител, но без промяна в нейния порядък, като същевременно нараства изчислителната сложност. Увеличаването на N (например, до $N = 16384$ или $N = 32768$, виж Фигура 2.2 и Таблица 2.2) не води до монотонно намаляване на грешката. Причината се свързва с повишената чувствителност на дробно-рационалната формула към малки

флуктуации в оценките на случайните величини $\theta^{(i)}$, както и с доминиращата роля на систематичната грешка.



Фигура 2.2: Сходимост към точната минимална собствена стойност на матрицата A при използване на РМК и РКМК при $k = 5, m = 5$ за различни на брой N

Практически резултатите показват, че за матрици с размерност от порядъка на няколко хиляди елемента, оптималният баланс се постига, когато m и k се избират според теоретичните оценки при предварително фиксирано ниво на систематичната грешка, а N е от същия порядък като размерността на матрицата. Прекалено малко N води до доминиране на стохастичната грешка, а прекомерно голямо N увеличава изчислителната цена без съществено подобрение на точността.

2.5 Обобщение

Оценяването на минималната собствена стойност λ_{min} на симетрична матрица чрез степенни стохастични алгоритми изисква изчисляване на обратната матрица A^{-1} , което е изчислително скъпо при големи размерности. Като ефективна алтернатива е изследван Стохастичен Резолвентен метод, който елиминира необходимостта от явно намиране на A^{-1} , използвайки вместо това резолвентната матрица $R_q = (\mathbf{I} - qA)^{-1}$ и нейните степени. Знакът на параметъра на ускорение q определя дали се оценява минималната или максималната собствена стойност.

Теоретичният анализ показва, че сходимостта на Резолвентния Степенен метод зависи съществено от параметъра q и степента m . Систематичната грешка намалява при увеличаване на m , а скоростта на сходимост се определя от отношението между крайните собствени стойности на матрицата. Прекъсването на биномния ред чрез параметъра k въвежда допълнителна систематична грешка, за която са получени аналитични оценки. Установена е линейна зависимост между m и k , позволяваща избор на минимални стойности при фиксирано ниво на систематична грешка и съответно намаляване на изчислителната сложност.

В стохастичната реализация възниква и стохастична грешка. Предложени са Резолвентен Монте Карло (РМК) и Резолвентен Квази-Монте Карло (РКМК) алгоритми, при които математическите очаквания се апроксимират чрез средните стойности от N реализации на случайните величини чрез прекъснатата верига на Марков с k прехода. При МК подхода грешката е от порядък $\mathcal{O}(N^{-1/2})$, докато при КМК се използва неравенството на Коксма-Хлавка и редици с малък дискрепанс (например, редица на Соболев), осигуряваща по-бърза сходимост. Общата точност се определя от баланса между систематичната грешка (поради прекъсване на реда) и стохастичната грешка (поради крайния брой реализации), което налага балансиран избор на параметрите m, k и N .

Получените теоретични оценки дават практически насоки за избор на m, k и N така, че двата вида грешки да бъдат съпоставими по порядък спрямо зададена точност. Числените експерименти за матрици с размерности $n = 500$ и $n = 1000$ потвърждават тези изводи. Най-добрата ефективност се постига при параметри, съответстващи на теоретично предсказания баланс, като увеличаването на N над определена граница не води до съществено подобрение. РКМК алгоритъм показва по-бърза сходимост спрямо РМК подхода, а хибридната GPU/CPU реализация позволява ефективно приложение при задачи с голяма размерност. Част от получените резултати, описани в тази глава, са публикувани в статиите [35, 36].

3. Стохастични методи за оценяване на пазарния риск на инвестиционни портфейли

В тази глава представяме приложение на разработения **Алгоритъм 1** към реален проблем от финансовата математика - оценяване на пазарния риск на инвестиционен портфейл. Целта е да се демонстрира практическата приложимост на стохастичните степенни алгоритми чрез анализ на спектралните характеристики на корелационната матрица на доходностите. Оценяването на екстремални собствени стойности позволява идентифициране на доминиращи рискови фактори и изследване на чувствителността на портфейла към пазарни колебания. По този начин теоретичните алгоритми се използват като инструмент за количествен анализ на риска и за подпомагане на инвестиционни решения.

3.1 Постановка на задачата

Основната задача, която си поставяме, е да се оцени максималната собствена стойност на корелационни матрици и да се анализира точността на получените оценки чрез прилагането на разработените *почти оптимални* Степенен Монте Карло (СМК) и Степенен Квази-Монте Карло (СКМК) алгоритми, представени в Глава 1. В рамките на СКМК алгоритъма се въвеждат допълнително два оператора: (1) оператор за пропускане (англ. *skipping operator*) и (2) оператор за прескачане през стъпка (англ. *leaping operator*). Тяхното използване е мотивирано от необходимостта да се намали влиянието на началните неравномерности, характерни за рандомизираните редици с малък дискрепанс.

3.2 Методи и данни

Финансовите данни, използвани за конструирането на корелационните матрици, са извлечени от платформите *Portfolio Visualizer* и *Yahoo Finance*, като достъпът до последната е осъществен чрез библиотеката *yFinance*. Оценяването на най-голямата собствена стойност на корелационните матрици се извършва с посочените алгоритми в Глава 1 и се анализира балансът между систематичната и стохастичната грешка. За интерпретация на резултатите в контекста на финансовия анализ се използва показателят, наречен *дял на обяснената вариация* (англ. *Fraction of Variance Explained*, FVE). Той измерва частта на вариацията, съответстваща на максималната собствена стойност и характеризира степента на влияние на доминиращия пазарен фактор.

3.3 Обща формулировка на ковариационна и корелационна матрица

Ковариацията и коефициентът на корелация са основни характеристики за описване на линейната зависимост между случайни величини. Ковариацията между две случайни величини X и Y се дефинира чрез:

$$\text{Cov}(X, Y) = E[(X - \mu_X)(Y - \mu_Y)], \quad (3.1)$$

където $\mu_X = E[X]$ и $\mu_Y = E[Y]$ [77, 91]. Тя измерва посоката на съвместната изменчивост, като може да приема както положителни, така и отрицателни стойности, но зависи от мащаба на променливите. Коефициентът на корелация представлява нормализирана форма на ковариацията и се задава чрез:

$$\text{Corr}(X, Y) = \rho_{XY} = \frac{\text{Cov}(X, Y)}{\sigma_X \sigma_Y}, \quad (3.2)$$

където σ_X и σ_Y са стандартните отклонения [77, 91]. Той приема стойности в интервала $[-1, 1]$ и позволява сравнение на линейните зависимости, независимо от мерните единици.

За случаен вектор $\mathbf{X} = (X_1, \dots, X_n)^T$ ковариационната матрица се дефинира като:

$$\Sigma = \text{Cov}(\mathbf{X}) = E[(\mathbf{X} - \mu)(\mathbf{X} - \mu)^T], \quad (3.3)$$

където елементите $\Sigma_{ij} = \text{Cov}(X_i, X_j)$ се пресмятат по формула (3.1), а диагоналните елементи $\Sigma_{ii} = \sigma_i^2$ представляват дисперсиите на съответните компоненти на случайния вектор. Съответно, корелационната

матрица е:

$$\hat{R} = [\rho_{ij}]_{i,j=1}^n, \quad \rho_{ij} = \text{Corr}(X_i, X_j), \quad i, j = 1, \dots, n, \quad (3.4)$$

като $\rho_{ii} = 1$ [42, 57].

В контекста на портфейлния анализ, корелационната матрица описва зависимостите между доходностите на активите, а най-голямата ѝ собствена стойност служи като индикатор за доминиращ пазарен фактор [42]. Високи стойности на тази характеристика са свързани със засилена пазарна синхронизация и ограничени възможности за диверсификация. Ковариационната матрица предоставя пълна информация за взаимозависимостите между активите и е фундаментален инструмент в задачите по оптимизация и управление на портфейли.

3.3.1 Конструирание на ковариационна и корелационна матрица от емпирични данни от финансови пазари

В емпиричните изследвания на финансовите пазари данните обикновено се представят чрез матрицата $B \in \mathbb{R}^{r \times n}$, където редовете r съответстват на наблюденията на доходностите във времето, а колоните n - на броя на различни активи. Всеки ред \mathbf{b}_i описва съвместното състояние на пазара в даден момент, което позволява анализ на взаимозависимостите между активите. Матрицата B се задава чрез:

$$B = [b_{ij}]_{i,j=1}^{r,n} = (\mathbf{b}_1, \dots, \mathbf{b}_i, \dots, \mathbf{b}_r)^T, \quad (3.5)$$

където $\mathbf{b}_i = (b_{i1}, \dots, b_{in})$ представлява n -мерен вектор от наблюдавани доходности за всички активи в момента i ($i = 1, \dots, r$).

Центриране на данните и извадкова ковариационна и корелационна матрица

Нека $\bar{\mathbf{b}}$ е n -мерният вектор на извадковите средни стойности на колоните на матрицата B . Центрираната матрица се дефинира по следния начин:

$$B_c = B - \mathbf{e}\bar{\mathbf{b}}^T, \quad (3.6)$$

където $\mathbf{e} \in \mathbb{R}^r$ е вектор от единици. Извадковата ковариационна матрица се определя чрез:

$$\Sigma = \frac{1}{r} B_c^T B_c. \quad (3.7)$$

При малки стойности на r във формула (3.7) делението се извършва с $r - 1$. Елементите Σ_{ij} измерват съвместната изменчивост между доходностите на активите, а диагоналните елементи Σ_{ii} съответстват на техните вариации. Рискът на портфейл Π с тегла \mathbf{w} се задава чрез квадратичната форма:

$$\text{Var}(\Pi) = \mathbf{w}^T \Sigma \mathbf{w}, \quad (3.8)$$

което подчертава ролята както на индивидуалните рискове, така и на взаимните зависимости между активите. Корелационната матрица се получава чрез нормализиране на ковариационната матрица със стандартните отклонения:

$$\hat{R} = D_\sigma^{-1/2} \Sigma D_\sigma^{-1/2}, \quad \rho_{ij} = \frac{\Sigma_{ij}}{\sigma_i \sigma_j}, \quad (3.9)$$

като $\rho_{ij} \in [-1, 1]$ и $\rho_{ii} = 1$. Тази матрица предоставя мащабно независима мярка за линейните зависимости между активите и е ключова за анализа на диверсификацията [42].

3.3.2 Дял на обяснената вариация

Дялът на обяснената вариация (FVE) е важна характеристика, свързана със спектралните свойства на ковариационната или корелационната матрица [47]. За симетрична матрица A със собствени стойности $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n \geq 0$, дялът на обяснената вариация (FVE), съответстващ на най-голямата собствена стойност λ_1 , се дефинира като:

$$FVE_1 = \frac{\lambda_1}{\sum_{j=1}^n \lambda_j} = \frac{\lambda_1}{\text{trace}(A)}, \quad (3.10)$$

а за първите k компоненти:

$$FVE_k = \frac{\sum_{i=1}^k \lambda_i}{\text{trace}(A)}. \quad (3.11)$$

В частния случай на корелационна матрица \hat{R} , за която $\text{trace}(\hat{R}) = n$, се получава:

$$FVE_1 = \frac{\lambda_1}{n}. \quad (3.12)$$

Тази величина измерва относителния принос на доминиращия фактор в общата вариация на системата. В контекста на финансовите пазари големите стойности на FVE_1 показват наличие на силен общ пазарен фактор и високо ниво на систематичен риск, което ограничава ефективността на диверсификацията. Обратно, по-малките стойности предполагат по-добра разпределеност на риска между независими фактори и по-благоприятни условия за оптимизация на портфейла. Следователно, оценяването на най-голямата собствена стойност на корелационната матрица и съответния дял на обяснената вариация представлява ключов етап при количествения анализ на пазарния риск.

3.3.3 Стохастични алгоритми за оценяване на максималната собствена стойност на корелационна матрица

При оценяването на най-голямата собствена стойност на разглежданите корелационни матрици се прилагат разработените в Глава 1 *почти оптимални* Степенен Монте Карло (СМК) и Степенен Квази-Монте Карло (СКМК) алгоритми. За тези алгоритми беше показано, че превъзхождат *класическите* им аналози, тъй като при съпоставими изчислителни разходи осигуряват по-висока числена точност и по-стабилна сходимост на оценките. *Почти оптималният* СКМК алгоритъм, означен като **Алгоритъм 1** в Глава 1, използва разбъркани редици с малък дискрепанс за конструиране на верига на Марков. Ако в стъпка 4 се генерират редици на Холтърн или Соболев, то те се реализират съответно чрез псевдокодовете **Алгоритъм 3** и **Алгоритъм 4**.

Algorithm 3 Pseudocode for generating a scrambled Halton Sequence with *skip* and *leap* parameters

- 1: CREATE FUNCTION: `generateHaltonSequence` ($N, k+1$)
 - 2: GENERATE: $p \leftarrow \text{haltonset}(k, \text{'skip'} = s_1, \text{'leap'} = l_1)$
 - 3: GENERATE: $p \leftarrow \text{scramble}(p, \text{'RR2'})$
 - 4: GENERATE: $\text{points} \leftarrow \text{net}(p, N)$
 - 5: OUTPUT: Return *points*
-

Algorithm 4 Pseudocode for generating a scrambled Sobol Sequence with *skip* and *leap* parameters

- 1: CREATE FUNCTION: `generateSobolSequence` ($N, k+1$)
 - 2: GENERATE: $p \leftarrow \text{sobolset}(k, \text{'skip'} = s_1, \text{'leap'} = l_1)$
 - 3: GENERATE: $p \leftarrow \text{scramble}(p, \text{'MatousekAffineOwen'})$
 - 4: GENERATE: $\text{points} \leftarrow \text{net}(p, N)$
 - 5: OUTPUT: Return *points*
-

При избор $s_1 = 0$ и $l_1 = 0$ се получават стандартни (*default*) версии на редиците съответно означени с $PQMC_{(H)}^{(d)}$ и $PQMC_{(S)}^{(d)}$. При ненулеви стойности на параметрите се прилагат операциите *skip* и *leap*, които позволяват контрол върху началните елементи и допълнително разбъркване на редиците, като съответните алгоритми се означават с $PQMC_{(H)}^{(s,l)}$ и $PQMC_{(S)}^{(s,l)}$.

За реализацията на *почти оптималния* СМК алгоритъм (**Алгоритъм 1a**) се използват генераторите на псевдослучайни числа *Mersenne Twister* (MT) и *Middle Square* (MS), като съответните варианти се означават с $PMC_{(MT)}$ и $PMC_{(MS)}$. Всички алгоритми са реализирани в среда MATLAB (версия 9.10, R2021a), като използваният синтаксис на функциите `haltonset`, `sobolset` и `scramble` в **Алгоритъм 3** и **Алгоритъм 4** е съобразен с официалната документация на MATLAB [93, 94].

3.4 Числени резултати и дискусия

В настоящата секция са представени числени резултати за оценяване на максималната собствена стойност на две корелационни матрици чрез *почти оптималните* СМК и СКМК алгоритми. Изследването е проведено върху два портфейла от активи с различна размерност, което позволява анализ на поведението на алгоритмите при различна степен на сложност и информационна наситеност на данните.

Първата корелационна матрица е конструирана на базата на месечните доходности на $n = 32$ глобални актива за периода от януари 2007 г. до март 2024 г., със съответни размерности $r = 207$ и $n = 32$. Втората

матрица е изградена от доходностите на $n = 490$ компании от индекса $S\&P500$ за периода 2021-2025 г., при $r = 5$ и $n = 490$. Чрез прилагане на центриране (т.е. формула (3.6)), оценка на ковариационната матрица (3.7) и нормализация (3.9) се получават корелационни матрици със съответни максимални собствени стойности и дялове на обяснена вариация:

- за $n = 32$: $\lambda_{max} = 18.14714$, $FVE_1 = 56.71\%$;
- за $n = 490$: $\lambda_{max} = 201.090512$, $FVE_1 = 41.04\%$.

Тези стойности са изчислени чрез вградените функции на МАТЛАВ и служат като референтни стойности при оценка на точността на алгоритмите.

В числените експерименти векторите \mathbf{h} и \mathbf{f} , използвани при конструирането на случайните величини $\theta^{(k)}$ и $\theta^{(k-1)}$, са избрани с L_1 -норма:

$$\|\mathbf{h}\| = \|\mathbf{f}\| = \sum_{i=1}^n \frac{1}{n} = 1,$$

което осигурява нормализация и стабилност на стохастичните оценки. За осигуряване на сходимост се разглежда матрицата $A = q\hat{R}$, където параметърът q се избира така, че да е изпълнено условието $\|A\| < 1$:

$$q = \frac{1}{|\lambda_1| + \delta}, \quad \delta > 0,$$

или при липса на предварителна оценка за λ_1 :

$$q = \frac{1}{\|\hat{R}\|_1 + \delta}, \quad \|\hat{R}\|_1 = \max_{1 \leq j \leq n} \sum_{i=1}^n |\rho_{ij}| \quad \text{и} \quad \delta > 0.$$

Този подход осигурява коректна приложимост на стохастичните методи и гарантира тяхната числена устойчивост при оценяването на максималната собствена стойност.

3.4.1 Случай (i): корелационна матрица с размерност $n = 32$

В Таблица 3.1 са представени числените резултати за оценяване на най-голямата собствена стойност на корелационната матрица \hat{R} при размерност $n = 32$. Получените числени резултати са реализирани чрез *почти оптималния* СКМ алгоритъм (**Алгоритъм 1a**). Използват се два генератора на псевдослучайни числа - *Mersenne Twister* (MT) и *Middle Square* (MS), съответно означени с $PMC_{(MT)}$ и $PMC_{(MS)}$. Оценяването на максималната собствена стойност $\lambda_1 = \lambda_{max}$ е извършено при различни стойности на броя преходи k във веригата на Марков и броя реализации N на случайните величини $\theta^{(k)}$ и $\theta^{(k-1)}$, като в Таблица 3.1 са представени и съответните абсолютни грешки. Анализът показва, че при фиксирано N увеличаването на k води до намаляване на грешката, което се стабилизира около $k = 8$ и се достига практическият оптимален баланс между систематичната и стохастичната компонента на грешката. От друга страна, с увеличаване на N се наблюдава монотонно намаляване на абсолютната грешка, което е в съответствие с класическите свойства на Монте Карло методите.

Минималните абсолютни грешки за двата алгоритъма са отбелязани в Таблица 3.1, като резултатите показват съпоставим порядък на точност при двата генератора. Визуализацията на Фигура 3.1 илюстрира сходимостта на оценките $\lambda_1^{(k)}$ към точната стойност λ_1 при фиксирано $k = 8$ и нарастващ брой реализации N . Наблюдава се, че и двата алгоритъма се стремят към референтната стойност, но $PMC_{(MT)}$ демонстрира по-висока точност и по-устойчива сходимост, особено при по-малки N . Практически, за постигане на точност до втория или третия знак след десетичната запетая са достатъчни $N = 512$ или $N = 1024$ симулации, което е напълно адекватно за приложения в портфейлния анализ.

Чрез използване на формула (3.12) за корелационна матрица с размерност $n = 32$ и оценките при $N = 1024$, се получават стойности $FVE_1 \approx 56.70\%$ за $PMC_{(MT)}$ и $FVE_1 \approx 56.72\%$ за $PMC_{(MS)}$. Това потвърждава, че приближенията на λ_1 с точност от порядъка $\mathcal{O}(10^{-3})$ позволяват надеждно изчисляване на дяла на обяснената вариация, който е ключов показател за оценка на систематичния риск и колективната динамика на активите. Следователно, *почти оптималният* СКМ алгоритъм представлява ефективен подход за анализ на корелационни матрици с неголеми размерности при относително малки изчислителни разходи.

В Таблица 3.2 са представени резултати от прилагането на *почти оптималния* СКМК алгоритъм за корелационната матрица ($n = 32$), реализиран чрез четири варианта: $PQMC_{(H)}^{(d)}$, $PQMC_{(S)}^{(d)}$, $PQMC_{(H)}^{(s,l)}$

и $PQMC_{(S)}^{(s,l)}$. Вариантите с индекс (d) съответстват при избор $s_1 = 0$ и $l_1 = 0$ в **Алгоритъм 3** и **Алгоритъм 4**, когато са използвани разбъркани редици на Холтън и Соболев с настройки по подразбиране. При ненулеви стойности на параметрите $skip$ и $leap$ се реализират вариантите (s, l) , които позволяват по-фин контрол върху равномерността на генерираните точки и потенциално подобряват сходимостта на оценките.

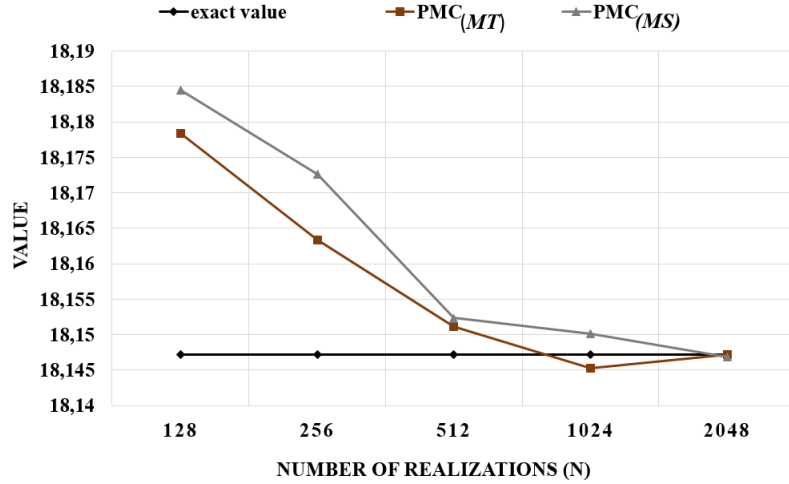
Таблица 3.1: Резултати за максималните стойности на $\lambda_{max}^{(k)}$ и абсолютната грешка за $PMC_{(MT)}$ и $PMC_{(MS)}$ при различни N и k , когато $\lambda_{max} = 18.14714$

| N | k | $\lambda_1^{(k)}$ using $PMC_{(MT)}$ | Absolute Error $PMC_{(MT)}$ | $\lambda_1^{(k)}$ using $PMC_{(MS)}$ | Absolute Error $PMC_{(MS)}$ |
|------|----------|--------------------------------------|-----------------------------|--------------------------------------|-----------------------------|
| 128 | 5 | 17.809541 | 0.3376 | 17.766635 | 0.3805 |
| | 6 | 17.955782 | 0.1914 | 17.931153 | 0.2160 |
| | 7 | 18.031908 | 0.1152 | 18.016774 | 0.1304 |
| | 8 | 18.178382 | 0.0312 | 18.184488 | 0.0373 |
| | 9 | 18.217756 | 0.0706 | 18.234182 | 0.0870 |
| | 10 | 18.259830 | 0.1127 | 18.278015 | 0.1309 |
| | 11 | 18.293246 | 0.1461 | 18.336586 | 0.1894 |
| 256 | 5 | 17.886432 | 0.2607 | 17.880110 | 0.2670 |
| | 6 | 18.013889 | 0.1333 | 18.025125 | 0.1220 |
| | 7 | 18.098671 | 0.0485 | 18.090287 | 0.0569 |
| | 8 | 18.163375 | 0.0162 | 18.172589 | 0.0254 |
| | 9 | 18.214595 | 0.0675 | 18.220052 | 0.0729 |
| | 10 | 18.265757 | 0.1186 | 18.270934 | 0.1238 |
| | 11 | 18.307841 | 0.1607 | 18.317397 | 0.1703 |
| 512 | 5 | 17.894326 | 0.2528 | 17.893265 | 0.2539 |
| | 6 | 18.019977 | 0.1272 | 18.014737 | 0.1324 |
| | 7 | 18.099165 | 0.0480 | 18.094732 | 0.0524 |
| | 8 | 18.151116 | 0.0040 | 18.152321 | 0.0052 |
| | 9 | 18.194868 | 0.0477 | 18.218433 | 0.0713 |
| | 10 | 18.231143 | 0.0840 | 18.244924 | 0.0978 |
| | 11 | 18.275571 | 0.1284 | 18.284115 | 0.1370 |
| 1024 | 5 | 17.977295 | 0.1698 | 17.947696 | 0.1994 |
| | 6 | 18.102537 | 0.0446 | 18.038992 | 0.1081 |
| | 7 | 18.118865 | 0.0283 | 18.098850 | 0.0483 |
| | 8 | 18.145221 | 0.0019 | 18.150097 | 0.0030 |
| | 9 | 18.176356 | 0.0292 | 18.214874 | 0.0677 |
| | 10 | 18.202606 | 0.0555 | 18.276008 | 0.1289 |
| | 11 | 18.250397 | 0.1033 | 18.294888 | 0.1477 |
| 2048 | 5 | 18.007899 | 0.1392 | 17.981049 | 0.1661 |
| | 6 | 18.106753 | 0.0404 | 18.056203 | 0.0909 |
| | 7 | 18.126731 | 0.0204 | 18.104529 | 0.0426 |
| | 8 | 18.147226 | 0.0001 | 18.146851 | 0.0003 |
| | 9 | 18.164272 | 0.0171 | 18.172111 | 0.0250 |
| | 10 | 18.193708 | 0.0466 | 18.215959 | 0.0688 |
| | 11 | 18.218186 | 0.0710 | 18.260127 | 0.1130 |

*Маркираните редове представляват оптималните k стойности за всяко N .

Числените резултати за *почти оптималните* СКМК алгоритми ($PQMC_{(H)}^{(s,l)}$ и $PQMC_{(S)}^{(s,l)}$) са получени при избор на параметри $s_1 = 1024$ и $l_1 = 128$ в **Алгоритъм 3** и **Алгоритъм 4**, което съответства на прилагане на операциите $skip$ и $leap$ при генерирането на разбърканите редици на Холтън и Соболев. Максималната собствена стойност $\lambda_1 = \lambda_{max}$ на корелационната матрица е оценена при различни стойности на броя преходи k и броя реализации N , като в Таблица 3.2 са представени и съответните абсолютни грешки. Получените резултати показват, че параметрите k и N оказват съществено влияние върху баланса между систематичната и стохастичната грешка, както и върху изчислителната сложност на алгоритмите. При фиксирано N увеличаването на k води до намаляване на абсолютната грешка, което се стабилизира около $k = 11$ за вариантите по подразбиране $PQMC_{(H)}^{(d)}$ и $PQMC_{(S)}^{(d)}$, и около $k = 13$ за вариантите с параметри $PQMC_{(H)}^{(s,l)}$ и $PQMC_{(S)}^{(s,l)}$, което показва достигане на практически оптимален баланс между компонентите на грешката. При фиксирано k и нарастващо N , алгоритмите $PQMC_{(H)}^{(d)}$ и $PQMC_{(S)}^{(d)}$ демонстрират монотонно намаляване на грешката и стабилна сходимост. За разлика от тях, при $PQMC_{(H)}^{(s,l)}$ и $PQMC_{(S)}^{(s,l)}$ се наблюдава значително подобрене при малки извадки ($N = 128, 256$), но при по-големи N грешката остава почти постоянна или дори нараства, което съответства на наблюденията в [52] относно

чувствителността към параметрите $skip$ и $leap$.



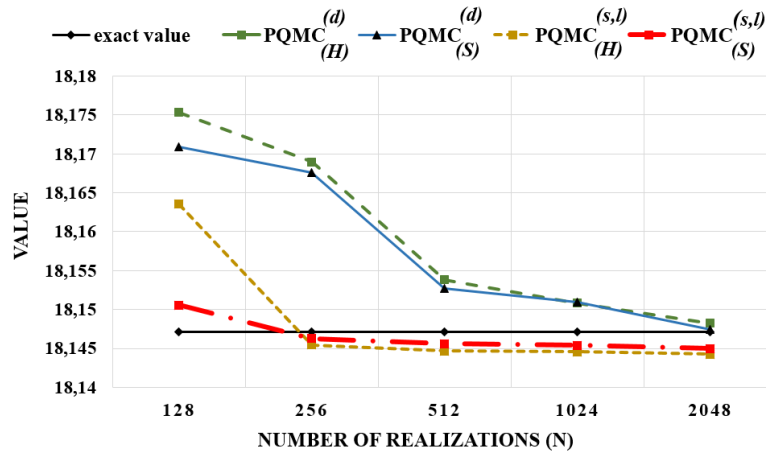
Фигура 3.1: Сравнение между точната максимална собствена стойност λ_1 и приближена собствена стойност $\lambda_1^{(k)}$ за $PMC_{(MT)}$, $PMC_{(MS)}$, получени в балансиращите случаи при различни стойности на N и $k = 8$

Таблица 3.2: Резултати за максимални собствени стойности за $\lambda_{max}^{(k)}$ и абсолютна грешка чрез сценарий по подразбиране и сценарий, използващ $skip$ и $leap$ при различни N и k , когато $\lambda_{max} = 18.14714$

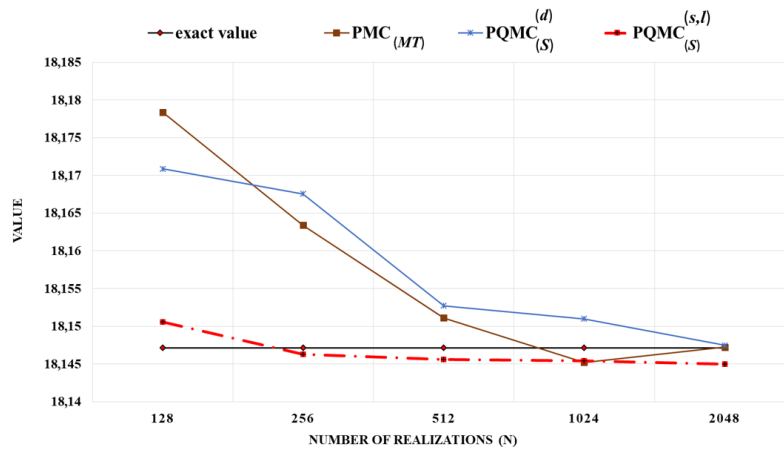
| N | k | $\lambda_1^{(k)}$ using $PQMC_{(H)}^{(d)}$ | Absolute Error $PQMC_{(H)}^{(d)}$ | $\lambda_1^{(k)}$ using $PQMC_{(S)}^{(d)}$ | Absolute Error $PQMC_{(S)}^{(d)}$ | $\lambda_1^{(k)}$ using $PQMC_{(H)}^{(s,l)}$ | Absolute Error $PQMC_{(H)}^{(s,l)}$ | $\lambda_1^{(k)}$ using $PQMC_{(S)}^{(s,l)}$ | Absolute Error $PQMC_{(S)}^{(s,l)}$ |
|------|-----------|--|-----------------------------------|--|-----------------------------------|--|-------------------------------------|--|-------------------------------------|
| 128 | 8 | 17.757547 | 0.3896 | 17.761379 | 0.3858 | 17.672733 | 0.4744 | 17.706637 | 0.4405 |
| | 9 | 17.888100 | 0.2590 | 17.888185 | 0.2590 | 17.787523 | 0.3596 | 17.792723 | 0.3544 |
| | 10 | 18.016669 | 0.1305 | 18.038214 | 0.1089 | 17.856741 | 0.2904 | 17.859089 | 0.2881 |
| | 11 | 18.175307 | 0.0282 | 18.170859 | 0.0237 | 17.903166 | 0.2440 | 17.914975 | 0.2322 |
| | 12 | 18.298811 | 0.1517 | 18.264832 | 0.1177 | 18.036970 | 0.1102 | 18.097564 | 0.0496 |
| | 13 | 18.479146 | 0.3320 | 18.474626 | 0.3275 | 18.145462 | 0.0017 | 18.146253 | 0.0009 |
| | 14 | 18.643362 | 0.4962 | 18.639031 | 0.4919 | 18.194240 | 0.0471 | 18.191902 | 0.0448 |
| 15 | 18.712889 | 0.5657 | 18.706247 | 0.5591 | 18.290732 | 0.1436 | 18.290924 | 0.1438 | |
| 256 | 8 | 17.882593 | 0.2645 | 17.883812 | 0.2633 | 17.761315 | 0.3858 | 17.763949 | 0.3832 |
| | 9 | 17.994408 | 0.1527 | 18.016913 | 0.1302 | 17.854804 | 0.2923 | 17.867712 | 0.2794 |
| | 10 | 18.061271 | 0.0859 | 18.062487 | 0.0847 | 17.936785 | 0.2104 | 17.921160 | 0.2260 |
| | 11 | 18.168944 | 0.0218 | 18.167556 | 0.0204 | 18.036010 | 0.1111 | 18.043959 | 0.1032 |
| | 12 | 18.284126 | 0.1370 | 18.251619 | 0.1045 | 18.097978 | 0.0492 | 18.098070 | 0.0491 |
| | 13 | 18.482373 | 0.3352 | 18.469322 | 0.3222 | 18.145462 | 0.0017 | 18.146253 | 0.0009 |
| | 14 | 18.545036 | 0.3979 | 18.516617 | 0.3695 | 18.209384 | 0.0622 | 18.185109 | 0.0380 |
| 15 | 18.608838 | 0.4617 | 18.607262 | 0.4601 | 18.289967 | 0.1428 | 18.270105 | 0.1230 | |
| 512 | 8 | 17.927397 | 0.2197 | 17.962654 | 0.1845 | 17.774098 | 0.3730 | 17.778972 | 0.3682 |
| | 9 | 18.020330 | 0.1268 | 18.021557 | 0.1256 | 17.846070 | 0.3011 | 17.867473 | 0.2797 |
| | 10 | 18.081105 | 0.0660 | 18.081198 | 0.0659 | 17.938077 | 0.2091 | 17.949995 | 0.1971 |
| | 11 | 18.153831 | 0.0067 | 18.152703 | 0.0056 | 18.009669 | 0.1375 | 18.014807 | 0.1323 |
| | 12 | 18.263301 | 0.1162 | 18.248667 | 0.1015 | 18.038078 | 0.1091 | 18.038550 | 0.1086 |
| | 13 | 18.425061 | 0.2779 | 18.411903 | 0.2648 | 18.144697 | 0.0024 | 18.145602 | 0.0015 |
| | 14 | 18.506259 | 0.3591 | 18.504204 | 0.3571 | 18.204814 | 0.0577 | 18.191490 | 0.0444 |
| 15 | 18.551586 | 0.4044 | 18.546687 | 0.3995 | 18.249184 | 0.1020 | 18.242947 | 0.0958 | |
| 1024 | 8 | 17.962214 | 0.1849 | 17.978111 | 0.1690 | 17.776881 | 0.3703 | 17.792702 | 0.3544 |
| | 9 | 18.035437 | 0.1117 | 18.042592 | 0.1045 | 17.867701 | 0.2794 | 17.887652 | 0.2595 |
| | 10 | 18.103422 | 0.0437 | 18.109582 | 0.0376 | 17.980078 | 0.1671 | 17.980502 | 0.1666 |
| | 11 | 18.150822 | 0.0037 | 18.150986 | 0.0038 | 18.010761 | 0.1364 | 18.037651 | 0.1095 |
| | 12 | 18.294487 | 0.1473 | 18.284447 | 0.1373 | 18.098692 | 0.0484 | 18.098718 | 0.0484 |
| | 13 | 18.385875 | 0.2387 | 18.384413 | 0.2373 | 18.144277 | 0.0029 | 18.144964 | 0.0022 |
| | 14 | 18.446470 | 0.2993 | 18.443825 | 0.2967 | 18.202841 | 0.0557 | 18.203207 | 0.0561 |
| 15 | 18.530500 | 0.3834 | 18.521017 | 0.3739 | 18.236167 | 0.0890 | 18.235788 | 0.0886 | |
| 2048 | 8 | 18.000431 | 0.1467 | 18.000619 | 0.1465 | 17.787254 | 0.3599 | 17.788274 | 0.3589 |
| | 9 | 18.061446 | 0.0857 | 18.064399 | 0.0827 | 17.879783 | 0.2674 | 17.887915 | 0.2592 |
| | 10 | 18.108451 | 0.0387 | 18.110469 | 0.0367 | 17.980078 | 0.1671 | 17.980502 | 0.1666 |
| | 11 | 18.148212 | 0.0011 | 18.147480 | 0.0003 | 18.010923 | 0.1362 | 18.040106 | 0.1070 |
| | 12 | 18.226453 | 0.0793 | 18.221684 | 0.0745 | 18.118819 | 0.0283 | 18.120939 | 0.0262 |
| | 13 | 18.385443 | 0.2383 | 18.379872 | 0.2327 | 18.144277 | 0.0029 | 18.144964 | 0.0022 |
| | 14 | 18.443228 | 0.2961 | 18.435573 | 0.2884 | 18.202381 | 0.0552 | 18.203555 | 0.0564 |
| 15 | 18.509033 | 0.3619 | 18.492101 | 0.3450 | 18.247648 | 0.1005 | 18.243565 | 0.0964 | |

*Маркираните редове представляват оптималните k стойности за всяко N .

Графичното сравнение на Фигура 3.2 потвърждава направените изводи, като илюстрира сходимостта на оценките $\lambda_1^{(k)}$ към точната стойност λ_1 при нарастващ брой реализации N . Алгоритмите $PQMC_{(S)}^{(d)}$ и $PQMC_{(H)}^{(d)}$ демонстрират устойчива и гладка сходимост, докато вариантите $PQMC_{(H)}^{(s,l)}$ и $PQMC_{(S)}^{(s,l)}$ показват по-добра точност при малки N , но ограничено подобрене при по-големи извадки. На Фигура 3.3 алгоритмите $PMC_{(MT)}$ и $PQMC_{(S)}^{(d)}$ се открояват с по-висока точност и стабилност при големи N , докато при малки извадки ($N \leq 512$) използването на СКМК алгоритми с параметри *skip* и *leap* е по-ефективно.



Фигура 3.2: $(PQMC_{(H)}^{(d)}, PQMC_{(S)}^{(d)})$ и $(PQMC_{(H)}^{(s,l)}, PQMC_{(S)}^{(s,l)})$. Сравнение между максималната собствена стойност λ_{\max} на корелационната матрица и оценените стойности $\lambda_1^{(k)}$, получени при $k = 11$ в случая по подразбиране и $k = 13$ в случая с пропускане и прескачане и различни стойности на N .



Фигура 3.3: Сравнение между максимална собствена стойност λ_{\max} и приближените собствени стойности $\lambda_1^{(k)}$, получени при $k = 8$ за $PMC_{(MT)}$, $k = 11$ за $PQMC_{(S)}^{(d)}$ и $k = 13$ за $PQMC_{(S)}^{(s,l)}$ при различни стойности на N .

Таблица 3.3: Приблизени стойности за дяла на обяснената вариация (FVE_1), получени при $k = 8$ за $PMC_{(MT)}$, $k = 11$ за $PQMC_{(S)}^{(d)}$ и $k = 13$ за $PQMC_{(S)}^{(s,l)}$ за корелационната матрица \hat{R} с размерност $n = 32$. Референтната стойност е $FVE_1 = 56.71\%$.

| N | $PMC_{(MT)}$ | $PQMC_{(S)}^{(d)}$ | $PQMC_{(S)}^{(s,l)}$ |
|------|--------------|--------------------|----------------------|
| 256 | 56.76% | 56.77% | 56.70% |
| 2048 | 56.71% | 56.71% | 56.70% |

В Таблица 3.3 са представени оценки за FVE_1 , получени чрез $PMC_{(MT)}$, $PQMC_{(S)}^{(d)}$ и $PQMC_{(S)}^{(s,l)}$, при

балансиращи стойности на k и $N = 256, 2048$, като референтната стойност е $FVE_1 = 56.71\%$. Резултатите показват, че при малки извадки $PQMC_{(S)}^{(s,l)}$ дава по-добри оценки, докато при големи N по-ефективни са $PMC_{(MT)}$ и $PQMC_{(S)}^{(d)}$, които постигат по-висока точност при по-малка изчислителна сложност. Високата стойност на FVE_1 за $n = 32$ показва наличие на доминиращ фактор, което означава по-висока концентрация на систематичен риск и ограничени възможности за диверсификация, тъй като активите реагират сходно на пазарни шокове.

3.4.2 Случай (ii): корелационна матрица с размерност $n = 490$

В Таблица 3.4 са представени числените резултати за оценяване на максималната собствена стойност $\lambda_1 = \lambda_{\max}$ на корелационна матрица \hat{R} с размерност $n = 490$, конструирана на базата на доходностите на компании от индекса $S\&P500$ за периода 2021–2025 г. (виж Приложение). Поради липса на пълни данни, десет компании са изключени от анализа. Оценяването е извършено чрез алгоритмите $PMC_{(MT)}$, $PQMC_{(H)}^{(d)}$ и $PQMC_{(S)}^{(d)}$, при различни стойности на броя преходи k и броя реализации N , като са анализирани и съответните абсолютни грешки. Оценяването е извършено чрез *почти оптималния* СКМ алгоритъм $PMC_{(MT)}$ и *почти оптималните* СКМК алгоритми ($PQMC_{(H)}^{(d)}$ и $PQMC_{(S)}^{(d)}$), използващи разбъркани редици на Холтън и Собол с настройки по *подразбиране*. Резултатите показват, че параметрите k и N оказват съществено влияние върху точността на оценките, като балансиращите стойности на броя преходи са $k = 8$ за $PMC_{(MT)}$ и $k = 11$ за $PQMC_{(H)}^{(d)}$ и $PQMC_{(S)}^{(d)}$, независимо от значително по-голямата размерност спрямо случая $n = 32$.

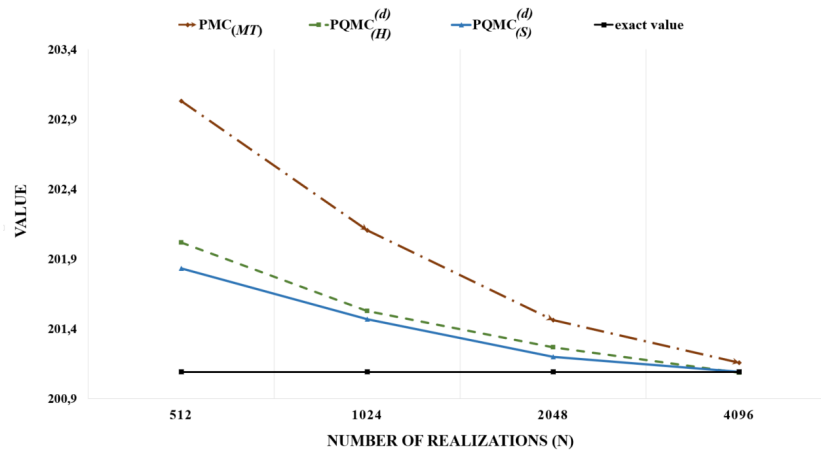
Таблица 3.4: Резултати за максимални собствени стойности за $\lambda_{\max}^{(k)}$ и абсолютна грешка при $PMC_{(MT)}$, $PQMC_{(H)}^{(d)}$ и $PQMC_{(S)}^{(d)}$ при различни N и k , когато $\lambda_{\max} = 201,090512$.

| N | k | $\lambda_1^{(k)}$ using $PMC_{(MT)}$ | Absolute Error $PMC_{(MT)}$ | k | $\lambda_1^{(k)}$ using $PQMC_{(H)}^{(d)}$ | Absolute Error $PQMC_{(H)}^{(d)}$ | $\lambda_1^{(k)}$ using $PQMC_{(S)}^{(d)}$ | Absolute Error $PQMC_{(S)}^{(d)}$ |
|------|----------|--------------------------------------|-----------------------------|-----------|--|-----------------------------------|--|-----------------------------------|
| 512 | 6 | 191.261222 | 9.8293 | 8 | 192.052259 | 9.0383 | 192.056796 | 9.0337 |
| | 7 | 197.764391 | 3.3261 | 9 | 193.305250 | 7.7853 | 194.4876692 | 6.6028 |
| | 8 | 203.030089 | 1.9396 | 10 | 196.545330 | 4.5452 | 196.561636 | 4.5289 |
| | 9 | 206.001210 | 4.9107 | 11 | 202.016871 | 0.9264 | 201.830255 | 0.7397 |
| | 10 | 209.202619 | 8.1121 | 12 | 207.216480 | 6.1260 | 206.994716 | 5.9042 |
| 1024 | 6 | 194.244715 | 6.8458 | 8 | 191.604757 | 9.4858 | 191.662089 | 9.4284 |
| | 7 | 199.671479 | 1.4190 | 9 | 194.355045 | 6.7355 | 194.443329 | 6.6472 |
| | 8 | 202.103664 | 1.0132 | 10 | 198.362863 | 2.7276 | 196.623242 | 4.4673 |
| | 9 | 204.706063 | 3.6156 | 11 | 201.526659 | 0.4361 | 201.467671 | 0.3772 |
| | 10 | 211.940453 | 10.8499 | 12 | 204.493635 | 3.4031 | 204.718275 | 3.6278 |
| 2048 | 6 | 196.531375 | 4.5591 | 8 | 191.522259 | 9.5683 | 191.737033 | 9.3535 |
| | 7 | 199.092487 | 1.9980 | 9 | 194.216871 | 6.8736 | 194.383618 | 6.7069 |
| | 8 | 201.462168 | 0.3717 | 10 | 198.552479 | 2.5380 | 198.866204 | 2.2243 |
| | 9 | 204.787620 | 3.6971 | 11 | 201.267671 | 0.1772 | 201.199260 | 0.1087 |
| | 10 | 206.264686 | 5.1942 | 12 | 207.590516 | 6.5000 | 206.670141 | 5.5796 |
| 4096 | 6 | 195.762474 | 5.3280 | 8 | 194.293635 | 6.7969 | 195.846272 | 6.2442 |
| | 7 | 197.402088 | 3.6884 | 9 | 198.498616 | 2.5919 | 198.498988 | 2.5915 |
| | 8 | 201.154640 | 0.0641 | 10 | 199.459470 | 1.6310 | 199.802438 | 1.2881 |
| | 9 | 202.887637 | 1.7971 | 11 | 201.083321 | 0.0072 | 201.091681 | 0.0012 |
| | 10 | 207.343376 | 6.2529 | 12 | 206.448996 | 5.3585 | 205.624451 | 4.5339 |

*Маркираните редове представляват оптималните k стойности за всяко N .

При фиксирано k се наблюдава монотонно намаляване на абсолютната грешка с увеличаване на броя реализации N , като минималните грешки са отбелязани в Таблица 3.4. Фигура 3.4 илюстрира, че всички алгоритми демонстрират устойчива сходност към референтната стойност λ_{\max} , като вариантите на СКМК, $PQMC_{(H)}^{(d)}$ и $PQMC_{(S)}^{(d)}$, показват по-бърза и по-гладка сходност спрямо $PMC_{(MT)}$. Получените оценки за λ_1 чрез $PQMC_{(H)}^{(d)}$ и $PQMC_{(S)}^{(d)}$ позволяват по-прецизно изчисляване на дяла на обяснената вариация FVE_1 чрез формула (3.12).

В Таблица 3.5 са представени съответните резултати, получени при балансиращите стойности $k = 8$ и $k = 11$, като референтната стойност е $FVE_1 = 41.04\%$. Данните показват, че при увеличаване на N използването на СКМК алгоритмите ($PQMC_{(H)}^{(d)}$ и $PQMC_{(S)}^{(d)}$) са за предпочитане пред $PMC_{(MT)}$ поради по-високата точност и при съпоставима изчислителна сложност.



Фигура 3.4: Сравнение между точната максимална собствена стойност (λ_{\max}) и максималната приближена собствена стойност ($\lambda_1^{(k)}$) за $PMC_{(MT)}$, $PQMC_{(H)}^{(d)}$, $PQMC_{(S)}^{(d)}$, получени когато $k = 8$ в случая за $PMC_{(MT)}$ и $k = 11$ в случая за $PQMC_{(*)}^{(d)}$ и различни стойности на N

Таблица 3.5: Приблизени стойности за дяла на обяснената вариация (FVE_1), получени при $k = 8$ за $PMC_{(MT)}$ и при $k = 11$ за $PQMC_{(H)}^{(d)}$ и $PQMC_{(S)}^{(d)}$ за корелационната матрица \hat{R} с размерност $n = 490$. Референтната стойност е $FVE_1 = 41.04\%$.

| N | $PMC_{(MT)}$ | $PQMC_{(H)}^{(d)}$ | $PQMC_{(S)}^{(d)}$ |
|------|--------------|--------------------|--------------------|
| 1024 | 41.25% | 41.13% | 41.12% |
| 2048 | 41.11% | 41.08% | 41.06% |
| 4096 | 41.05% | 41.04% | 41.04% |

Стойността $FVE_1 \approx 41\%$ за корелационната матрица с размерност $n = 490$ показва, че вариацията не е концентрирана в първата главна компонента, а е разпределена между множество фактори. Това свидетелства за по-сложна структура на зависимостите между активите и предполага по-висок потенциал за диверсификация. В същото време, този резултат налага използването на многофакторни модели за адекватно описание на пазарния риск, тъй като еднофакторните подходи биха били недостатъчни за улавяне на наблюдаваната динамика.

3.5 Обобщение

Получените резултати потвърждават практическата приложимост и числената ефективност на *почти оптималните* стохастични алгоритми за оценяване на максималната собствена стойност на корелационни матрици при анализ на пазарния риск. Изследваните матрици, с размерности $n = 32$ и $n = 490$, са конструирани на базата на данни от *Portfolio Visualizer* и *Yahoo Finance* (виж Приложение). Анализът на корелационните матрици с размерности $n = 32$ и $n = 490$ показва, че максималната собствена стойност λ_1 има ясна интерпретация като индикатор за наличието и интензитета на доминиращ общ пазарен фактор. Чрез съответния дял на обяснената вариация FVE_1 се получава количествена оценка както за нивото на систематичния риск, така и за ограниченията пред ефективната диверсификация на портфейла.

Числените експерименти ($n = 32$) показват, че при *почти оптималните* СКМК алгоритми балансът между стохастичната и систематичната грешка се постига при $k = 8$. При *почти оптималните* СКМК алгоритми с разбъркани редици на Собол и Холтън с настройки по *подразбиране* този баланс се достига при $k = 11$, а при използване на операторите *skip* и *leap* - при $k = 13$. Освен това, с увеличаване на броя реализации N се наблюдава монотонно намаляване на абсолютната грешка при СКМК и СКМК алгоритмите с настройки по *подразбиране*, което потвърждава тяхната стабилна сходимост.

При оценяване на максималната собствена стойност на матрица с размерност ($n = 490$) *почти оптималните* СКМК алгоритми с настройки по *подразбиране* демонстрират по-добра сходимост и по-висока точност спрямо *почти оптималните* СКМК алгоритми. Използването на операторите *skip* и *leap* е ефективно предимно при малък брой симулации N , като при по-големи извадки тяхното предимство намалява и не гарантира допълнително подобрение на точността. Част от представените резултати са публикувани

в [38].

В обобщение, разработените стохастични степенни алгоритми могат да бъдат разглеждани като надежден подход за числено оценяване на спектрални характеристики на корелационни матрици. Това от своя страна позволява задълбочен анализ на систематичния риск, колективната динамика на активите и ефективността на диверсификацията. В бъдещи изследвания тези алгоритми могат да бъдат прилагани в задачи, свързани с финансово моделиране, управление на портфейли и подпомагане на инвестиционни решения в условия на нарастваща размерност.

Заклучение

Настоящата дисертация е посветена на разработването, анализа и приложението на стохастични числени методи от типа МК и рандомизирани КМК методи за оценяване на екстремални собствени стойности на симетрични матрици и използването им при практически задачи, свързани с оценка на риска на инвестиционни портфейли. Основната цел е изграждането на стабилна и изчислително ефективна методологична рамка, която осигурява управляем баланс между систематичната и стохастичната грешка и е приложима при задачи с висока размерност.

В рамките на изследването са обосновани и изследвани Степенният (Резолвентният) Монте Карло метод и рандомизираният Степенен (Резолвентен) Квази-Монте Карло метод за оценяване на екстремални собствени стойности. Съществен резултат е прецизирането на конструкцията на случайните величини при използването на почти оптимална преходна плътност във веригата на Марков, както и идентифицирането на класове симетрични матрици, при които тази конструкция се опростява. На тази основа са разработени почти оптимални алгоритми (Степенен и Резолвентен МК и КМК), описани чрез псевдокод, за които са дадени оценки за изчислителната сложност и условия за балансиране на статистическата и систематичната грешка. Теоретичните аргументи и експерименталните резултати потвърждават, че предложените почти оптимални конструкции водят до намаляване на дисперсията и повишаване на точността при същия или по-малък изчислителен разход спрямо класическите стохастични аналози.

Експерименталната верификация е проведена както с генератори на псевдослучайни числа (MT и MS), така и с редици с малък дискрепанс (редици на Соболев и Холтън) и техни разбъркани варианти. Числените изследвания показват измерим ефект от избора на генератор и типа редица върху точността и стабилността на оценките, като при разглежданите тестове генераторът на псевдослучайни числа MT демонстрира предимство спрямо генератора MS, а редиците на Соболев, спрямо редиците на Холтън. Установено е също, че за постигане на сравним баланс на грешките рандомизираният Степенен (Резолвентен) КМК алгоритъм може да изисква малко по-голям брой преходи във веригата на Марков, което се компенсира от по-добрата сходимост при ограничени изчислителни ресурси.

Разработените алгоритми са приложени към задача от финансовата математика, свързана с управлението на риска на инвестиционен портфейл. Максималната собствена стойност на корелационната матрица на активи е използвана като индикатор за концентрация на риска, като оценяването ѝ е реализирано чрез предложените СКМ и СКМК алгоритми. При реализиране на СКМК алгоритмите с разбъркани редици с малък дискрепанс (редици на Соболев и Холтън) са изследвани варианти и с използване на операторите *skip* и *leap*.

В обобщение, дисертационният труд демонстрира, че прилагането на разработените алгоритми за получаване на спектрални оценки с целенасочено балансиране на грешките води до по-висока точност, стабилност и изчислителна ефективност. Очертани са и перспективи за бъдещо развитие, включително разширяване към устойчиви (робастни) модели, динамични ковариационни структури и допълнителни изчислителни оптимизации за разреждени симетрични матрици с големи размерности от порядъка на милиони.

Списък на публикациите по дисертацията

За написването на дисертацията са използвани следните публикации:

1. **Gurova, S.-M.**, Karaivanova, A., *Quasi-Monte Carlo Algorithms for Eigenvalue Problems, 4th International Conference Numerical Methods for Scientific Computations and Advanced Applications, NMSCAA'24*, 18-22, (2024) (short communication)
<https://parallel.bas.bg/~elilkova/nmscaa24.pdf>
2. **Gurova, S.-M.**, Atanassov, E., Karaivanova, A., *A Resolvent Quasi-Monte Carlo Method for Estimating the Minimum Eigenvalues Using the Error Balancing*, In: Lirkov, I., Margenov, S. (eds), *Large-Scale Scientific Computing, LSSC'23, LNCS*, vol. **13952**, 394-403, (2024), (SJR:0.352, Q2, 2024)
https://doi.org/10.1007/978-3-031-56208-2_40
3. **Gurova, S.-M.**, Gurov, T., Karaivanova, A., *On the Stochastic Power Algorithms for Estimating the Maximum Eigenvalue of Symmetric Matrices*, **Contemp. Math.**, Universal Wiser Publisher Pte. Ltd, **6(5):7223-7249**, (2025) doi.org/10.37256/cm.6520257781 (IF:2.5, 2025, WoS: Q1, <https://wos-journal.info/journalid/23187>)

4. Gurova, S.-M., Gurov, T., Karaivanova, A., *Eigenvalue Estimation in Portfolio Risk: The Role of Skipping and Leaping in Sobol and Halton Sequences*, In: Lirkov, I., Margenov, S. (eds), Large-Scale Scientific Computing, LSSC'25, LNCS, 16061,(accepted)

Апробация на резултатите

Резултатите, които са включени в дисертацията, са докладвани на семинари на секцията "Скалируеми алгоритми и киберсигурност с Център по високопроизводителни пресмятания" към ИИКТ-БАН, както и на международни конференции. Изнесени са следните доклади, които са свързани с темата на дисертацията:

1. Gurova, S.-M., Monte Carlo Method for Estimating Eigenvalues Using Error Balancing, The 13th International Conference on Large-Scale Scientific Computations, LSSC'21, 07-11.06.2021, Sozopol, Bulgaria. <https://parallel.bas.bg/Conferences/SciCom21/program.html>
2. Gurova, S.-M., A Quasi-Monte Carlo method for estimation of eigenvalues using error balancing, The 15th International Conference on Monte Carlo and Quasi-Monte Carlo methods in Scientific Computing, MCQMC'22, 17-22.07.2022, Linz, Austria. https://www.ricam.oeaw.ac.at/events/conferences/mcqm2022/schedule/MCQMC2022_book_final_version.pdf
3. Gurova, S.-M., Power Monte Carlo method using randomized low discrepancy sequences, the 17th Annual Meeting of the Bulgarian Section of SIAM, 20-22.12.2022, BGSIAM'22, Sofia, Bulgaria. http://www.math.bas.bg/bgsiam/docs/bgsiam_2022_program.pdf
4. Gurova, S.-M., A Resolvent Quasi-Monte Carlo Method for Estimating the Minimum Eigenvalues Using the Error Balancing, The 14th International Conference on Large-Scale Scientific Computations, LSSC'23, 05-09.06.2023, Sozopol, Bulgaria. <https://parallel.bas.bg/~elilkova/lssc23.pdf#page=17>
5. Gurova, S.-M., Quasi-Monte Carlo Algorithms for Eigenvalue Problems, Fourth International Conference Numerical Methods for Scientific Computations and Advanced Applications, NMSCAA'24, 17-21.06.2024, Sozopol, Bulgaria. <https://parallel.bas.bg/Conferences/nmscaa24/program.html>

Статиите представят резултати, получени в рамките на един международен и три национални научни проекта, като единият от тях е разделен на две фази, а именно:

1. Международен научен проект

- Нови скалируеми алгоритми с използване на Монте Карло и Квази-Монте Карло методи и тяхното приложение в областта на машинното обучение, финансиран от CAF America, 2021-2024, ръководител: проф. д-р Емануил Атанасов.

2. Национални научни проекти

- "Център за върхови постижения по Информатика и информационни и комуникационни технологии" (фаза 1) по договор BG05M2OP001-1.001-0003, финансиран от оперативна програма "Наука и образование за интелигентен растеж" 2014-2020, съфинансиран от Европейския съюз чрез Европейските структурни и инвестиционни фондове, ръководител: чл.-кор. д.м.н. Светозар Маргенов.
- "Национален геоинформационен център (НГИЦ) за мониторинг, оценка и прогнозиране на природни и антропогенни рискове и бедствия" обект от "Националната пътна карта за научна инфраструктура 2020-2027" (договори: Д01-321/30.11.2023, Д01-164/28.07.2022), финансиран от Министерство на образованието и науката, ръководител: доц. д-р Петя Трифонова.
- "Национален център за високопроизводителни и разпределени пресмятания" (НЦВРП) обект от "Националната пътна карта за научна инфраструктура 2020-2027" (договори: Д01-98/26.06.2025, Д01-325/01.12.2023), финансиран от Министерство на образованието и науката, ръководител: проф. д-р Анета Караиванова.
- "Център за върхови постижения в областта на информатиката и информационните и комуникационни технологии" (фаза 2) по договор BG16RFPR002-1.014-0018-C01, финансиран по програмата за научни изследвания, иновации и дигитализация за интелигентна трансформация 2021-2027 и съфинансиран от Европейския съюз, ръководител: чл.-кор. д.м.н. Светозар Маргенов.

Основни научни и научно-приложни приноси

Получените научни и научно-приложни приноси напълно съответстват на поставените цели и задачи на дисертационния труд. Те пряко отразяват обосноваването, разработването и експерименталното изследване на Степенен МК/КМК алгоритми и Резолвентен МК/КМК алгоритми за оценка на екстремни собствени стойности на симетрични матрици. Освен това, приносите са потвърдени и в реална задача от финансовата математика чрез оценяване на пазарния риск на портфейл от глобални активи.

1. Основните научни приноси на настоящата дисертация са:

- Обосновани и изследвани са Степенният (Резолвентният) Монте Карло метод и рандомизираният Степенен (Резолвентен) Квази-Монте Карло метод за оценяване на екстремални собствени стойности на симетрични матрици. Прецизирана е конструкцията на случайните величини $\theta^{(k)}$ и $\theta^{(k-1)}$ при използване на почти оптимална преходна плътност във веригата на Марков. Формулирани са класове симетрични матрици, за които се опростява конструкцията на случайните величини в стохастичните степенни алгоритми. Получени са теоретични оценки за зависимостта между степента на резолвентната матрица и дължината на веригата на Марков в зависимост от ускоряващия параметър и нормата на матрицата.
- Разработени са почти оптимални СМК и СКМК алгоритми, както и почти оптимални РМК и РКМК алгоритми, базирани на предложената прецизирана конструкция на случайните величини, които са представени чрез псевдокод. Дадени са оценки за изчислителната сложност на алгоритмите и условията за баланс между стохастичната и систематична грешка. Аргументирано е, че предложените алгоритми водят до намаляване на дисперсията спрямо класическите аналози.
- Предложена е практическа задача от областта на финансовата математика за оценяване на пазарния риск на инвестиционни портфейли. Конструираната корелационна матрица за съответния портфейл е използвана за оценяването на максималната собствена стойност, която е индикатор за концентрация на пазарния риск. Оценяването се извършва чрез разработените в Глава 1 почти оптимален Степенен МК и Степенен КМК алгоритми, като допълнително Степенният КМК алгоритъм е модифициран чрез използване на параметрите *skip* и *leap*.

2. Основните научно-приложни приноси на настоящата дисертация са:

- Предложените почти оптимални СМК алгоритми са реализирани с МТ и MS генератори на псевдослучайни числа, а СКМК алгоритмите, с разбъркани редици на Соболев и Холтън. Числените експерименти за оценяване на максималната собствена стойност на плътни симетрични матрици с големи размерности демонстрират превъзходството на почти оптималните алгоритми над класическите им аналози както по точност, така и по изчислителна сложност.
- Числено е изследван балансът между стохастичната и систематичната грешка, който зависи съществено от броя на преходите k във веригата на Марков и от броя реализации N на случайните величини $\theta^{(k)}$ и $\theta^{(k-1)}$. Резултатите показват, че СКМК алгоритмите се нуждаят от 2-3 стъпки повече във веригата на Марков в сравнение със СМК алгоритмите. Изборът на генератор на псевдослучайни числа или на разбърканата редица с малък дискрепанс има измерим ефект върху точността и стабилността. В разгледаните експерименти МТ е предпочитан пред MS, а редиците на Соболев пред редиците на Холтън.
- Почти оптималният РМК алгоритъм е реализиран с МТ генератор на псевдослучайни числа, а почти оптималният РКМК алгоритъм с разбъркана редица на Соболев чрез BRODA's Sobol RSG генератор. Числените експерименти, проведени на НРС кълъстер с NVIDIA Tesla V100 (32 GB) GPU карти, за оценяване на минималната собствена стойност, показват, че балансът между стохастичната и систематичната грешка зависи от броя на преходите k във веригата на Марков, броя реализации N на случайните величини $\theta^{(k)}$ и $\theta^{(k-1)}$, степента m на резолвентната матрица и параметъра за ускорение q . Резултатите, получени с рандомизирана редица на Соболев и с МТ генератор за оценяване на минималната собствена стойност на симетрични матрици, демонстрират един и същ порядък. От съществено значение е възможността да се намери константа, която отделя минималната собствена стойност от нулата.
- Проведени са числени експерименти за оценяване на максималната собствена стойност на две корелационни матрици, която е индикатор за концентрация на пазарния риск. Първата корелационна матрица е конструирана въз основа на доходите на 32 глобални актива, докато

при втората са използвани доходите на 490 компании от *S&P500*, за конкретни времеви периоди (виж Приложение). Оценяването е извършено чрез почти оптималния СКК алгоритъм с МТ/MS генератори на псевдослучайни числа и чрез почти оптималния СКМК алгоритъм, както без допълнителни настройки (*default* вариант), така и с използване на параметрите *skip* и *leap* при разбъркани редици на Соболев и Холтън.

Приложение

За провеждането на числените експерименти са използвани генерирани симетрични матрици (Глава 1, Глава 2) и реални финансови данни (Глава 3). Симетричните матрици с размерности 100 и 500, използвани за числените експерименти в Глава 1, са генерирани в софтуерната среда MATLAB, докато тези с размерности 500 и 1000 от Глава 2 са част от алгоритмичната реализация, разработена на програмния език C++.

Финансовите данни, използвани за конструирането на корелационната матрица с размерност $n = 32$ в Глава 3, са извлечени от уеб платформата *Portfolio Visualizer* (www.portfoliovisualizer.com). Изследваният времеви обхват обхваща периода от януари 2007 г. до март 2024 г., като са използвани реални пазарни показатели. Въз основа на тази емпирична информация е конструиран инвестиционен портфейл. Портфейлът се състои от тридесет и два глобални актива, класифицирани според техния инвестиционен тип, както следва:

1. **Осемнадесет акции** - „US Stock Market“, „US Large Cap Value“, „US Large Cap Growth“, „US Mid Cap“, „US Mid Cap Value“, „US Mid Cap Growth“, „US Small Cap Value“, „US Small Cap Growth“, „US Large Cap“, „US Small Cap“, „US Micro Cap“, „Global ex-US Stock Market“, „European Stocks“, „Pacific Stocks“, „Emerging Markets“, „Intl Developed ex-US Market“, „International ex-US Small Cap“, „International ex-US Value“;
2. **Девет облигации** - „Short Term Treasury“, „Intermediate Term Treasury“, „10-year Treasury“, „Long Term Treasury“, „Total US Bond Market“, „Global Bonds (USD Hedged)“, „Short-Term Investment Grade“, „TIPS (Treasury Inflation-Protected Securities)“, „Corporate Bonds“;
3. **Четири алтернативни инвестиции** - „REIT (Real Estate Investment Trusts)“, „Gold“, „Precious Metals“, „Commodities“;
4. **Един паричен актив и парични еквиваленти** - „Cash“.

Финансовите данни, използвани за конструирането на корелационната матрица с размерност $n = 490$ в Глава 3, са доходите на 490 компании от *S&P500 (Standard&Poor's 500)*, индекс за оценка на състоянието на фондовия пазар в САЩ, за 5-годишен период от 2021 до 2025 година. Данните за тяхната доходност са изтеглени от *Yahoo Finance* (<https://finance.yahoo.com/>) посредством програмната библиотека *yFinance*, достъпна чрез хранилището на *PyPI* (<https://pypi.org/project/yfinance/>). Чрез тази библиотека се осъществява автоматизирано извличане на исторически ценови серии и пазарни индикатори чрез интерфейс към уеб страницата на *Yahoo Finance*, което позволява програмно изтегляне на данни за конкретни финансови инструменти в зададен времеви интервал и последващата им обработка в аналитична среда.

1. Генерираните симетрични матрици с размерности 100, 500, използвани за числените експерименти в Глава 1, са публикувани в платформата **GitHub** и са достъпни на следния адрес:
 - https://github.com/smgurova/Symmetric-matrices_chapter-1.git
2. Финансовият отчет, предоставен от *Portfolio Visualizer*, е наличен в платформата **GitHub** на следния адрес:
 - <https://github.com/smgurova/AssetAllocation.git>
3. Корелационната матрица на тридесетте и два глобални актива, използвана за числените експерименти в Глава 3, е достъпна в платформата **GitHub** на следния адрес:
 - https://github.com/smgurova/correlation-matrix_chapter-4.git
4. Данните за доходността на 490 компании от *S&P500 (Standard&Poor's 500)* за времеви период от 2021 до 2025 година и образуваната от тях корелационна матрица, използвана за числените експерименти в Глава 3, са достъпни в платформата **GitHub** на следния адрес:
 - <https://github.com/smgurova/Data-S-P500.git>

Библиография

- [1] Alexandrov, V., Atanassov, E., Dimov, I., Branford, S., Thandavan, A., Weihrauch, C., Parallel Hybrid Monte Carlo Algorithms for Matrix Computations. In: Sunderam, V.S., van Albada, G.D., Sloot, P.M.A., Dongarra, J. (eds.) Computational Science – ICCS 2005. LNCS, vol. 3516, Springer, 752-759, (2005), https://doi.org/10.1007/11428862_102
- [2] Alexandrov, V., Davila, D., Esquivel-Flores, O., Karaivanova, A., Gurov, T., Atanassov, E., On Monte Carlo and quasi-Monte Carlo Methods for Matrix Computations, Lecture Notes in Computer Science, **10665**, Springer-Verlag, 249-257, (2018), DOI: https://doi.org/10.1007/978-3-319-73441-5_26
- [3] Alexandrov, V., Esquivel-Flores, O., Ivanovska, S., Karaivanova, A., On the Preconditioned Quasi-Monte Carlo Algorithm for Matrix Computations. In: Lirkov, I., Margenov, S., Waśniewski, J. (eds.) Large-Scale Scientific Computing. LSSC 2015. Lecture Notes in Computer Science, vol. **9374**, Springer, Cham., 163-171, (2015), https://doi.org/10.1007/978-3-319-26520-9_17
- [4] Alexandrov, V., Esquivel-Flores, O.A., Towards Monte Carlo preconditioning approach and hybrid Monte Carlo algorithms for matrix computations, Computers & Mathematics with Applications, vol. **70**(11), 2709-2718, (2015), <https://doi.org/10.1016/j.camwa.2015.08.035>
- [5] Almulihi A.H. H., High-Performance Quasi-Monte Carlo Integration and Applications, Dissertations. 3452, Western Michigan University, (2019), <https://scholarworks.wmich.edu/dissertations/3452>, (Last accessed: January 2026)
- [6] Branford, S., Hybrid Monte Carlo Methods for Linear Algebra Problems, Ph.D. dissertation, School of Systems Engineering, The University of Reading, (2009)
- [7] BRODA's Sobol RSG: broda.co.uk/software.html (Last accessed: 2026/01/11)
- [8] Brouwer, A.E., Haemers, W.H., Spectra of Graphs, Springer, (2012) <https://doi.org/10.1007/978-1-4614-1939-6>
- [9] Caffisch, R. E., Monte Carlo and quasi-Monte Carlo methods, Acta Numerica, Cambridge University Press, 1-49, (1998) <https://www.math.ucla.edu/~caflisch/Pubs/Pubs1995-1999/actaNumerica1998.pdf> (Last accessed: 2026/01/11)
- [10] Chi, H. Scrambled Quasirandom Sequences and Their Applications, Florida State University, (2004) http://pur1.flvc.org/fsu/fd/FSU_migr_etd-3823, (Last accessed: 2026/01/11)
- [11] Chi, H, Beerli, P., Evans, D. W., Mascagni, M., On the Scrambled Sobol Sequence, LNCS 3516, Springer, 775-782, (2005)
- [12] Chung, F., Spectral Graph Theory, CBMS Regional Conference Series in Mathematics 92, AMS, (1997).
- [13] Dick, J., Pillichshammer, F., Digital Nets and Sequences: Discrepancy Theory and Quasi-Monte Carlo Integration, Cambridge University Press (2010), <https://doi.org/10.1017/CB09780511761188>
- [14] Dekking, M., A Modern Introduction to Probability and Statistics. Springer, 181–190, (2005), ISBN: 9781852338961
- [15] Dimov, I., Monte Carlo algorithms for linear problems, Pliska Stud. Math. Bulgar., **13**, 57-77, (2000)
- [16] Dimov, I., Dimov, T., Gurov, T., A new iterative Monte Carlo approach for inverse matrix problem, Journal of Computational and Applied Mathematics, **92**(1), 15-35, (1998)
- [17] Dimov, I., Tonev, O., Monte Carlo algorithms: performance analysis for some computer architectures, Journal of Computational and Applied Mathematics, **48**, 253-277, (1993)
- [18] Dimov, I., Alexandrov, V., Karaivanova, A., Implementation of Monte Carlo Algorithms for Eigenvalue Problem Using MPI, PVM/MPI'98, LNCS, **1497**, 346-353, (1998)
- [19] Dimov, I., Philippe, B., Karaivanova, A., Weihrauch, C., Robustness and applicability of Markov chain Monte Carlo algorithms for eigenvalue problems, Applied Mathematical Modelling, **32**(8), 1511-1529, (2008)

- [20] Dimov, I, Alexandrov, V., Karaivanova, A., Parallel resolvent Monte Carlo algorithms for linear algebra problems, *Mathematics and Computers in Simulation*, **55**(1), 25-35, Published by Elsevier BV., (2005)
- [21] Dimov, I., Karaivanova, A.: A Power Method with Monte Carlo Iterations, *Proceedings Recent Advances in Numerical Methods and Applications II*, World Scientific, 239-247, (1999), https://doi.org/10.1142/9789814291071_0022
- [22] Dimov, I., *Monte Carlo Methods for Applied Scientists*, World Scientific Publishing Co. Pte. Ltd., (2007), ISBN: 13 978-981-02-2329-8
- [23] Dimov, I., Karaivanova, A., Iterative Monte Carlo algorithms for linear algebra problems, *Springer Lecture Notes in Computer Science*, **1196**, 66-77, (1996)
- [24] Dimov, I., Karaivanova, A., Parallel computations of eigenvalues based on a Monte Carlo approach, *Monte Carlo Methods and Applications*, Vol. **4**, Num. 1: 33-52, 1998
- [25] Dimov, I.T., et al.: Monte Carlo Algorithms for Calculating Eigenvalues. In: Niederreiter, H., et.al. (eds.) *Monte Carlo and Quasi-Monte Carlo Methods*, 1996, *Lecture Notes in Statistics*, vol. **127**, Springer, New York, NY., (1998)
- [26] Erkan, U., A precise and stable machine learning algorithm: eigenvalue classification (EigenClass), *Neural Comput & Applic*, **33**, 5381-5392, (2021), <https://doi.org/10.1007/s00521-020-05343-2>
- [27] Fathi, B., Liu, B., Alexandrov, V., Mixed MC parallel algorithms for Matrix Computation, *LNCS*, **2330**, 609-618, (2002)
- [28] Faure H., Lemieux Ch.: Generalized Halton sequences in 2008: A comparative study. *ACM Transactions on Modeling and Comp. Sim*, Vol. **19**, Issue 4, 1-31, (2009), <https://doi.org/10.1145/1596519.1596520>
- [29] Faure, H., Discr pance de suites associ es a un syst me de num ration (en dimension s), *Acta Arithmetica*, Vol. **41**, 337-351, (1982)
- [30] Faure, H., Good permutations for extreme discrepancy, *Journal of Number Theory*, Vol. **42**, 47-56, (1992)
- [31] Fischer, H., *A History of the Central Limit Theorem: From Classical to Modern Probability Theory*, Sources and Studies in the History of Mathematics and Physical Sciences. New York: Springer, (2011), <http://dx.doi.org/10.1007/978-0-387-87857-7>
- [32] Fishman, G.S., Multiplicative congruential random number generators, *Math. Comp.*, **54**, 331-344, (1990)
- [33] Golub, G., Loan, Ch. F. Van, *Matrix computation*, 4th edition, The Johns Hopkins University Press, (2013), ISBN: 13-978-1-4214-0794-4, <https://math.ecnu.edu.cn/~jypan/Teaching/books/2013%20Matrix%20Computations%204th.pdf>
- [34] Gilboa, G., *Nonlinear Eigenproblems in Image Processing and Computer Vision*, Publisher: Springer Cham, ed.number 1, (2018), eBook ISBN: 978-3-319-75847-3
- [35] Gurova, S.-M., Atanassov, E., Karaivanova, A., A Resolvent Quasi-Monte Carlo Method for Estimating the Minimum Eigenvalues Using the Error Balancing, In: Lirkov, I., Margenov, S. (eds.), *Large-Scale Scientific Computing, LSSC 2023*, *Lecture Notes in Computer Science*, vol. **13952**, 394-403, (2024)
- [36] Gurova, S.-M., Karaivanova, A., Quasi-Monte Carlo Algorithms for Eigenvalue Problems, *Fourth International Conference Numerical Methods for Scientific Computations and Advanced Applications, NMSCAA'24*, 17-21, (2024) (short communication) <https://parallel.bas.bg/~elilkova/nmscaa24.pdf>
- [37] Gurova, S.-M., Gurov, T., Karaivanova, A., On the Stochastic Power Algorithms for Estimating the Maximum Eigenvalue of Symmetric Matrices, *Contemp. Math.*, Universal Wiser Publisher Pte. Ltd., **6**(5): 7223-7249, (2025) <https://doi.org/10.37256/cm.6520257781>
- [38] Gurova, S.-M., Gurov, T., Karaivanova, A., Eigenvalue Estimation in Portfolio Risk: The Role of Skipping and Leaping in Sobol and Halton Sequences, In: Lirkov, I., Margenov, S. (eds), *Large-Scale Scientific Computing, LSSC'25*, *LNCS*, **16061**,(accepted for publication)
- [39] Halton, J., On the efficiency of certain quasi-random sequences of points in evaluating multi-dimensional integrals, *Numerische Mathematik*, **2**, 84-90, (1960), <https://doi.org/10.1007/BF01386213>

- [40] Halton, J., Smith, GB, Algorithm 247: Radical-Inverse Quasi-Random Point Sequence, Communications of the ACM, Vol. **7**, 701-702, (1964)
- [41] Hammersley, J.M., Handscomb, D.C., Random, Pseudorandom, and Quasirandom Numbers. In: Monte Carlo Methods. Monographs on Applied Probability and Statistics. Springer, Dordrecht, (1964), https://doi.org/10.1007/978-94-009-5819-7_3
- [42] Higham, N. J., Computing the nearest correlation matrix—a problem from finance, IMA Journal of Numerical Analysis, **22**(3): 329–343, (2002), <https://doi.org/10.1093/imanum/22.3.329>
- [43] Isaacson, E., Keller, H. B., Analysis of Numerical Methods, Dover Publications, Inc., N.Y., Wiley, (1996), ISBN: 0-486-68029-0
- [44] Ivanovska, S., Karaivanova, A., Manev, N., Numerical Integration Using Sequences Generating Permutations. In: Lirkov, I., Margenov, S., Waśniewski, J. (eds) Large-Scale Scientific Computing. LSSC 2011. Lecture Notes in Computer Science, vol. **7116**, Springer, Berlin, Heidelberg, 455-463, (2011), https://doi.org/10.1007/978-3-642-29843-1_51
- [45] Joe, S., Sloan, I., Implementation of a lattice method for numerical multiple integration, ACM Transactions on Mathematical Software (TOMS), Vol. **19**, Issue 4, 523–545, (1993), <https://doi.org/10.1145/168173.168425>
- [46] Joe S., Kuo F.Y.: Constructing Sobol Sequences with Better Two-Dimensional Projections, SIAM Journal on Scientific Computing, **30**(5): 2635-2654, (2008), <https://doi.org/10.1137/070709359>
- [47] Johnstone, I. M., On the distribution of the largest eigenvalue in principal components analysis, Ann. Statist. **29**(2), 295-327, (2001), <https://doi.org/10.1214/aos/1009210544>
- [48] Karaivanova, A., Quasi-Monte Carlo Methods for Some Linear Algebra Problems, Convergence and complexity, Serdica J. Computing, **4**, 57-72, (2010)
- [49] Kolda, T. and Bader, B., Tensor Decompositions and Applications, SIAM Review, **51**(3), 455-500, (2009) <https://doi.org/10.1137/07070111X>
- [50] Kroese, D.P., Taimre, T., Botev, Z. I., Handbook of Monte Carlo Methods, John Wiley & Sons Inc., New York, N.Y., (2011), ISBN:9780470177938
- [51] Laloux, L., Cizeau, P., Bouchaud, J.-P., Potters, M., Noise Dressing of Financial Correlation Matrices, Phys. Rev. Lett., **83**(7), 1467-1470, American Physical Society, (1999), <https://doi.org/10.1103/PhysRevLett.83.1467>
- [52] Lemieux, C., Monte Carlo and Quasi-Monte Carlo Sampling. Springer, (2009), <https://doi.org/10.1007/978-0-387-78165-5>
- [53] L'Ecuyer, P. Uniform random number generation. Ann Oper Res, **53**, 77–120, (1994), <https://doi.org/10.1007/BF02136827>
- [54] L'Ecuyer, P., Efficient and portable combined random number generators, Communications of the ACM, vol. **31**, Issue 6, 742-751, (1988), <https://doi.org/10.1145/62959.62969>
- [55] Lobo, P.D.C., Mikhailov, M., Ozisik, M.N., On the complex eigenvalues of Luikov system of equations, Drying Technology, **5**(2), 273-286, (1987), <https://doi.org/10.1080/07373938708916540>
- [56] Loperfido, N., Skewness-based projection pursuit: A computational approach, Computational Statistics and Data Analysis, **120**, 42-57, (2018), <https://doi.org/10.1016/j.csda.2017.11.001>
- [57] Macdonald, R.R., Correlation and Covariance Matrices. In Wiley StatsRef: Statistics Reference Online (eds. N. Balakrishnan, T. Colton, B. Everitt, W. Piegorisch, F. Ruggeri and J.L. Teugels), (2014), <https://doi.org/10.1002/9781118445112.stat06481>
- [58] Mascagni, M., Karaivanova, A., A Parallel Quasi-Monte Carlo Method for Solving Systems of Linear Equations, P.M.A. Sloot et al. (Eds.): ICCS 2002, LNCS **2330**, Springer-Verlag Berlin Heidelberg, 598-608, (2002)
- [59] Mascagni, M., Karaivanova, A., A Parallel Quasi-Monte Carlo Method for Computing Extremal Eigenvalues, K.-T. Fang et al. (eds.), Monte Carlo and Quasi-Monte Carlo Methods, Springer-Verlag Berlin Heidelberg, 369 – 380, (2000)

- [60] Mascagni, M., Karaivanova, A., Matrix Computations Using Quasirandom Sequences, Lecture Notes in Computer Science, (Wulkov, Yalamov, Wasniewsky eds.), Vol. **1988**, Springer, 552-559, (2001)
- [61] Mascagni, M., Karaivanova, A., A Monte Carlo Approach for Finding More than One Eigenpair. In: Dimov, I., Lirkov, I., Margenov, S., Zlatev, Z. (eds) Numerical Methods and Applications - NMA 2002, Lecture Notes in Computer Science, vol.**2542**, Springer, 123-131, (2003), https://doi.org/10.1007/3-540-36487-0_13
- [62] Matsumoto, M., Nishimura, T., Mersenne Twister: A 623-dimensionally Equidistributed Uniform Pseudo-Random Number Generator, ACM Transactions on Modeling and Computer Simulation, vol. 8(1), 3-30, (1998), <https://doi.org/10.1145/272991.272995>
- [63] Mehrdoust, F., Algorithm for Estimating the Condition Number of Matrices, Math. Reports, **15(65)**(3), 203-210, (2013)
- [64] Muminov M.I., Rasulov T.H.: Embedded Eigenvalues of a Hamiltonian in Bosonic Fock Space. Commun. Math. Anal., **17**(1), 1 - 22, (2014)
- [65] Mushtaq, A., Noreen, A., Olaussen, K., Numerical Solutions of Quantum Mechanical Eigenvalue Problems, Front. Phys., **8**:390, (2020), <https://doi.org/10.3389/fphy.2020.00390>
- [66] Niederreiter, H., Random Number Generation and Quasi-Monte Carlo Methods, SIAM (1992), <https://doi.org/10.1137/1.9781611970081>
- [67] Owen, A., Randomly Permuted (t,m,s)-Nets and (t, s)-Sequences. In: Niederreiter, H., Shiue, P.JS. (eds.) Monte Carlo and Quasi-Monte Carlo Methods in Scientific Computing. Lecture Notes in Statistics, vol. **106**, Springer, New York, 299-317, (1995), https://doi.org/10.1007/978-1-4612-2552-2_19
- [68] Owen, A., Monte Carlo variance of scrambled net quadrature, SIAM Journal on Numerical Analysis, Vol. **34**, No. 5, 1884-1910, (1997), <https://www.jstor.org/stable/2952022>
- [69] Owen, A., Variance with alternative scramblings of digital nets, ACM Transactions on Modeling and Computer Simulations, **13**(4), 363-378, (2003)
- [70] Padányi, V., Herendi, T., Generalized Middle-Square Method, Annales Mathematicae et Informaticae, **56**, 95-108, (2022), <https://doi.org/10.33039/ami.2022.12.003>
- [71] Park, Stephen K.; Miller, Keith W., Random Number Generators: Good Ones Are Hard To Find, Communications of the ACM, vol. **31**(10), 1192-1201, (1988), doi:10.1145/63039.63042
- [72] Rometsch, M., Quasi-Monte Carlo Methods in Finance with Application to Optimal Asset Allocation, Druck Diplomatica Verlag GmbH, Hamburg (2008).
- [73] Rotter, I., A non-Hermitian Hamilton operator and the physics of open quantum systems, J. Phys. A: Math. Theor., Vol. **42**, No. 15, 153001, (2009), <https://doi.org/10.1088/1751-8113/42/15/153001>
- [74] Sobol I., Shukhman B., Integration with quasirandom sequences: Numerical experience., Internat. J. Modern Phys. C, **6**(2), 263-275 (1995)
- [75] Sobol I., Asotsky D., Kreinin A., Kucherenko S., Construction and Comparison of High-Dimensional Sobol' Generator, WILMOTT magazine, 64-79, (2012)
- [76] Sobol I., Uniformly distributed sequences with an additional property of uniformity, USSR Computational Mathematics and Mathematical Physics, **16**(5), 236-242, (1976), [https://doi.org/10.1016/0041-5553\(76\)90154-3](https://doi.org/10.1016/0041-5553(76)90154-3)
- [77] Spiegel, M. R., Theory and Problems of Probability and Statistics, 2nd ed. New York: McGraw-Hill, (1992).
- [78] Straßburg J, Alexandrov, V., Enhancing Monte Carlo preconditioning methods for matrix computations, Procedia Computer Science, **29**, 1580-1589, (2014), <https://doi.org/10.1016/j.procs.2014.05.143>
- [79] Straßburg, J., On hybrid and resilient Monte Carlo methods for linear algebra problems, Ph.D. dissertation, The University of Reading, (2014)
- [80] Sturmfels, B., Tensors and Their Eigenvectors, Notices of the AMS, **63**(6), 604-606, (2016) <http://dx.doi.org/10.1090/noti1389>

- [81] Sugita, H., Monte Carlo Method, Random Number, and Pseudorandom Number, MSJ Memoirs, **25**: 133, (2011), <https://doi.org/10.2969/msjmemoirs/025010000>
- [82] Taft K., Vajargah, F., MC method for solving system of linear algebraic equations with minimum markov chains, International Conference PDPTA'2000, Las Vegas, (2000)
- [83] Vajargah F., et al., Parallel Monte Carlo algorithms for matrix computations, Ph.D. dissertation, The University of Reading, (2001)
- [84] Vajargah, F., Hassanzadeh, Z., Improvements on the hybrid Monte Carlo algorithms for matrix computations, Sadhana, Indian Academy of Science, **44**(1), <https://doi.org/10.1007/s12046-018-0983-y>
- [85] Van der Corput, J., Verteilungsfunktionen (Erste Mitteilung), Proceedings of the Koninklijke Akademie van Wetenschappen te Amsterdam (in German), **38**: 813–821, (1935), Zbl 0012.34705
- [86] Wang, X., Hickernell, F., Randomized Halton sequences, Mathematical and Computer Modelling, **32**(7–8), 887–899, (2000) [https://doi.org/10.1016/S0895-7177\(00\)00178-3](https://doi.org/10.1016/S0895-7177(00)00178-3)
- [87] Боянов, Б., Лекции по Числени методи, изд. Дарба София, (2008), ISBN: 954-90126-1
- [88] Димитров, Б., Янев, Н., Вероятности и статистика, Софттех, София, (2007)
- [89] Караиванова, А., Стохастични числени методи и симулации, изд. Деметра ЕООД, (2012), ISBN: 978-954-9526-78-3
- [90] Сендов, Б., Попов, В., Числени методи, Университетско издание „Св. Климент Охридски“, първа част, (1996), ISSN: 954-07-0994-6
- [91] Стоянов, Й., Миразчийски, И., Игнатов, Ц., Танушев, М., Ръководство за упражнения по теория на вероятностите. Издателство „Наука и изкуство“, София, (1976).
- [92] Соболев, И. М., Численные методы Монте-Карло, изд. Москва, Наука, (1973)
- [93] MathWorks: haltonset - Create a Halton point set. <https://www.mathworks.com/help/stats/haltonset.html>, (Last accessed: 2026/01/11)
- [94] MathWorks: sobolset - Create a Sobol point set. <https://www.mathworks.com/help/stats/sobolset.html>, (Last accessed: 2026/01/11)