



**ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΕΦΑΡΜΟΣΜΕΝΩΝ ΜΑΘΗΜΑΤΙΚΩΝ ΚΑΙ
ΦΥΣΙΚΩΝ ΕΠΙΣΤΗΜΩΝ**

ΔΙΑΤΜΗΜΑΤΙΚΟ ΠΡΟΓΡΑΜΜΑ ΜΕΤΑΠΤΥΧΙΑΚΩΝ ΣΠΟΥΔΩΝ

**«ΜΑΘΗΜΑΤΙΚΗ ΠΡΟΤΥΠΟΠΟΙΗΣΗ σε ΣΥΓΧΡΟΝΕΣ ΤΕΧΝΟΛΟΓΙΕΣ
και στα ΧΡΗΜΑΤΟΟΙΚΟΝΟΜΙΚΑ»**

**ΣΤΟΧΑΣΤΙΚΟΙ ΕΚΤΙΜΗΤΕΣ ΤΟΥ ΙΧΝΟΥΣ
ΣΥΝΑΡΤΗΣΕΩΝ ΠΙΝΑΚΩΝ**

ΠΑΡΑΣΚΕΥΗ ΦΙΚΑ

ΑΘΗΝΑ, ΜΑΡΤΙΟΣ 2016

ΠΑΡΑΣΚΕΥΗ ΦΙΚΑ
ΑΡΙΘΜΟΣ ΜΗΤΡΩΟΥ 09312043

Στοχαστικοί Εκτιμητές του Ύψους
Συναρτήσεων Πινάκων

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ ΓΙΑ ΤΟ Π.Μ.Σ.
ΜΑΘΗΜΑΤΙΚΗ ΠΡΟΤΥΠΟΠΟΙΗΣΗ ΣΕ ΣΥΓΧΡΟΝΕΣ
ΤΕΧΝΟΛΟΓΙΕΣ ΚΑΙ ΣΤΑ ΧΡΗΜΑΤΟΟΙΚΟΝΟΜΙΚΑ

ΕΠΙΒΛΕΠΩΝ ΚΑΘΗΓΗΤΗΣ
ΧΡΗΣΤΟΣ ΚΟΥΚΟΥΒΙΝΟΣ, ΚΑΘΗΓΗΤΗΣ ΕΜΠ

ΕΠΙΤΡΟΠΗ

- *Χρήστος Κουκουβίνος*
Καθηγητής ΕΜΠ (επιβλέπων)

- *Φιλία Βόντα*
Αναπ. Καθηγήτρια ΕΜΠ

- *Ιωάννης Σπηλιώτης*
Αναπ. Καθηγητής ΕΜΠ

Θα ήθελα να εκφράσω τις ευχαριστίες μου στον επιβλέπων
καθηγητή μου, κ. Χρήστο Κουκουβίνο, για τις γνώσεις που μου
μετέδωσε πάνω σε θέματα της Στατιστικής κατά τη φοίτησή μου
στο μεταπτυχιακό πρόγραμμα.

Περίληψη

Στην διπλωματική αυτή εργασία αναπτύσσονται στοχαστικοί εκτιμητές του ίχνους συναρτήσεων πινάκων, που συμβολίζεται ως $\text{Tr}(f(A))$, για κατάλληλη συνάρτηση f και συμμετρικό πίνακα A , ώστε να αποφεύγεται ο απ' ευθείας υπολογισμός του πίνακα $f(A)$.

Υπάρχουν πολλές εφαρμογές στις οποίες απαιτείται η εκτίμηση του ίχνους μιας συνάρτησης πίνακα, όπως η φυσική, η στατιστική, η ανάλυση δικτύων κ.α. Ο υπολογισμός του πίνακα $f(A)$ μπορεί να έχει υψηλό υπολογιστικό κόστος, ιδιαίτερα σε μεγάλους πίνακες που απαιτούνται στις εφαρμογές. Επομένως, κρίνεται αναγκαία η ύπαρξη εκτιμητών. Οι στοχαστικοί εκτιμητές που αναπτύσσονται στηρίζονται στον υπολογισμό της ποσότητας $\frac{1}{N} \sum_1^N x_i^T f(A) x_i$ για διαφορετικά διανύσματα x_i , όπου N είναι το μέγεθος του δείγματος (το πλήθος των διανυσμάτων). Αποδεικνύεται ότι η αναμενόμενη τιμή της ποσότητας $x^T f(A) x$, για κατάλληλο διάνυσμα x και συμμετρικό πίνακα A , ισούται με το ίχνος του πίνακα $f(A)$.

Επιπλέον, καθώς οι εκτιμητές που παράγονται είναι στατιστικής φύσης, είναι αναμενόμενο ότι η μελέτη και η εφαρμογή στατιστικών σχεδιασμών μπορεί να βελτιώσει την ποιότητα και την απόδοση τους στην προσέγγιση του ίχνους $\text{Tr}(f(A))$. Διαφορετικοί στατιστικοί σχεδιασμοί θα συγκριθούν και αριθμητικά παραδείγματα από πραγματικές εφαρμογές θα συζητηθούν.

Λέξεις Κλειδιά: στοχαστικός εκτιμητής, ίχνος, συνάρτηση πίνακα

Abstract

In this postgraduate work, stochastic estimates for the trace of functions of matrices, denoted as $\text{Tr}(f(A))$, are developed, where f is a suitable function and A a symmetric matrix. This aims to the avoidance of the explicit computation of the matrix $f(A)$.

The trace of a matrix function arises in many disciplines such as in physics, in statistics, in network analysis, etc. The computation of a function of the matrix A can have a high computational cost, especially for large matrices that are required in applications. Therefore, it is necessary to estimate the trace of $f(A)$. Stochastic estimates have been proposed, which are based on the computation of the quantity $\frac{1}{N} \sum_1^N x_i^T f(A) x_i$ for different sample vectors x_i , where N is the sample size (the number of the sample vectors). It can be proved that the expected value of the quantity $x^T f(A) x$, for suitable vector x and symmetric matrix A , equals to the $\text{Tr}(f(A))$.

Moreover, since the proposed estimates are of a statistical nature, it is natural to expect that the application of some statistical designs may improve their quality and their effectiveness in estimating the trace $\text{Tr}(f(A))$. Various statistical designs will be compared and numerical examples stemming from real applications will be discussed.

keywords: stochastic estimator, trace, matrix function

Περιεχόμενα

Εισαγωγή 13

1 Στοχαστικοί εκτιμητές του ίχνους 25

1.1 Το θεώρημα του Hutchinson 25

1.2 Hutchinson εκτιμητής του ίχνους $\text{Tr}(f(A))$ 32

1.3 Άλλοι στοχαστικοί εκτιμητές 33

1.4 Υπολογιστική πολυπλοκότητα - συμπεράσματα 36

1.5 Διαστήματα Εμπιστοσύνης 37

2 Στοχαστικοί εκτιμητές του ίχνους μέσω πινάκων Hadamard 41

2.1 Εκτιμητές μέσω της μεθόδου του Hutchinson 43

2.2 Δειγματοληψία από Sylvester Hadamard πίνακες 46

2.3 Δειγματοληψία από Hadamard πίνακες μέσω doubling construction 52

2.4 Εκτιμήσεις χρησιμοποιώντας άλλους σχεδιασμούς 53

3 Αριθμητικά παραδείγματα 59

Βιβλιογραφία 69

Εισαγωγή

Στο εισαγωγικό αυτό κεφάλαιο, θα παραθέσουμε κάποιες βασικές έννοιες και ιδιότητες οι οποίες θα χρειαστούν στα επόμενα κεφάλαια.

Φασματική Ανάλυση

Το επόμενο θεώρημα περιγράφει τη φασματική ανάλυση του A σε περίπτωση που ο A είναι συμμετρικός πίνακας (βλ. [22], [6]).

Θεώρημα 1. Φασματικό Θεώρημα (Spectral Theorem)

Εστω $A \in \mathbb{R}^{p \times p}$ συμμετρικός πίνακας. Υπάρχει ορθοκανονικός πίνακας V του οποίου οι στήλες v_k είναι τα κανονικοποιημένα ιδιοδιανύσματα του A και ένας διαγώνιος πίνακας Λ του οποίου τα διαγώνια στοιχεία είναι οι ιδιοτιμές λ_k του A τέτοιοι ώστε

$$(1) \quad A = V\Lambda V^T$$

με $VV^T = I_p$, $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_p)$ και $\lambda_1 \geq \dots \geq \lambda_p$.

Ορισμός 1.

Μία συνάρτηση f λέμε ότι ορίζεται στο φάσμα του συμμετρικού πίνακα $A \in \mathbb{R}^{p \times p}$ αν υπάρχουν οι τιμές $f(\lambda_i)$, $i = 1, \dots, p$ [35].

Ορισμός 2.

Εστω f μία συνάρτηση που ορίζεται στο φάσμα του συμμετρικού πίνακα $A \in \mathbb{R}^{p \times p}$ και έχει τη φασματική ανάλυση (1). Τότε $f(A) = Vf(\Lambda)V^T$.

Επομένως, για συμμετρικούς πίνακες, ο $f(A)$ έχει τα ίδια ιδιοδιανύσματα με τον πίνακα A και οι ιδιοτιμές του δίνονται εφαρμόζοντας την f στις ιδιοτιμές του πίνακα A [38], [66].

Πόρισμα 1.

Αν ο πίνακας A είναι συμμετρικός τότε και πίνακας $f(A)$ είναι συμμετρικός.

Απόδειξη

Προκύπτει άμεσα αν λάβουμε υπόψιν μας την φασματική ανάλυση του πίνακα $f(A) = Vf(\Lambda)V^T$. Παρατηρούμε ότι $f(A)^T = (Vf(\Lambda)V^T)^T = Vf(\Lambda)^T V^T = Vf(\Lambda)V^T = f(A)$.

Έστω x ένα τυχαίο, μη μηδενικό διάνυσμα στο \mathbb{R}^p και $V = [v_1, \dots, v_p]$

Από την φασματική ανάλυση του πίνακα A προκύπτει ότι,

$$Ax = \sum_{k=1}^p \lambda_k(v_k, x)v_k$$

και

$$f(A)x = \sum_{k=1}^p f(\lambda_k)(v_k, x)v_k.$$

Επομένως,

$$x^T f(A)x = \sum_{k=1}^p f(\lambda_k)(v_k, x)^2.$$

Ορισμός 3.

Έστω A ένας πραγματικός τετραγωνικός πίνακας διάστασης p ($A \in \mathbb{R}^{p \times p}$).

Ορίζεται ίχνος (*trace*) του πίνακα A το άθροισμα των διαγώνιων στοιχείων του, δηλαδή,

$$\text{Tr}(A) = a_{11} + a_{22} + \dots + a_{pp} = \sum_{i=1}^p a_{ii},$$

όπου a_{ii} είναι η είσοδος που βρίσκεται στην i -οστή γραμμή και στην i -οστή στήλη του A .

Λήμμα 1.

Το ίχνος ενός πίνακα είναι το άθροισμα όλων των ιδιοτιμών του.

Απόδειξη

Για το χαρακτηριστικό πολυώνυμο του πίνακα A έχουμε

$$\xi_A(x) = (a_{11} - x)(a_{22} - x) \dots (a_{pp} - x) + \delta(x),$$

όπου $\delta(x)$ είναι ένα πολυώνυμο βαθμού το πολύ $p - 2$. Ο συντελεστής του x^{p-1} στο δεξί μέλος είναι $(-1)^{p-1}(a_{11} + a_{22} + \dots + a_{pp}) = (-1)^{p-1}\text{Tr}(A)$.

Το $\xi_A(x)$ έχει p ρίζες, που είναι οι ιδιοτιμές του πίνακα A , έστω $\lambda_1, \dots, \lambda_p$. Επομένως, θα έχουμε

$$\xi_A(x) = (-1)^p(x - \lambda_1)(x - \lambda_2) \dots (x - \lambda_p).$$

Σε αυτήν την έκφραση του χαρακτηριστικού πολυωνύμου, ο συντελεστής του x^{p-1} στο δεξί μέλος είναι $(-1)^{p-1}(\lambda_1 + \dots + \lambda_p)$.

Επομένως, άμεσα προκύπτει ότι $\text{Tr}(A) = \lambda_1 + \dots + \lambda_p$. ■

Για το ίχνος του πίνακα $f(A)$ ισχύει

$$\text{Tr}(f(A)) = \sum_{k=1}^p f(\lambda_k).$$

Λήμμα 2. (Ιδιότητες του ίχνους)

Έστω $A, B \in \mathbb{R}^{p \times p}$ και σταθερά $c \in \mathbb{R}$. Ισχύουν οι ακόλουθες ιδιότητες για το ίχνος.

- Είναι γραμμική συνάρτηση.

$$\text{Tr}(A + B) = \text{Tr}(A) + \text{Tr}(B),$$

$$\text{Tr}(cA) = c\text{Tr}(A).$$

- $Tr(AB) = Tr(BA)$ (ακόμα και αν $AB \neq BA$)
- $Tr(A) = Tr(A^T)$

Σε πολλές εφαρμογές απαιτείται το ίχνος συνάρτησης ενός πίνακα A , επομένως καθίσταται αναγκαία η ανάπτυξη τέτοιων εκτιμητών, καθώς σε ένας τέτοιος υπολογισμός μπορεί να έχει υψηλό υπολογιστικό κόστος, ιδιαίτερα για μεγάλους πίνακες που απαιτούνται σε πραγματικές εφαρμογές. Οι στοχαστικοί εκτιμητές βασίζονται στον υπολογισμό της ποσότητας $\frac{1}{N} \sum_{i=1}^N x_i^T f(A) x_i$ για συγκεκριμένα διανύσματα x_i , όπου N είναι το μέγεθος του δείγματος (το πλήθος των διανυσμάτων). Η αναμενόμενη τιμή της ποσότητας $x^T f(A) x$ ισούται με το ίχνος του πίνακα $f(A)$.

Είναι επίσης σημαντική η ανάπτυξη μεθόδων ώστε να αποφευχθεί ο απ' ευθείας υπολογισμός των ποσοτήτων $x^T f(A) x$, που χρειάζονται στον υπολογισμό των εκτιμητών του ίχνους του αντιστρόφου πίνακα $f(A)$, $\frac{1}{N} \sum_{i=1}^N x_i^T f(A) x_i$.

Για την συνάρτηση $f(A) = A^r$, $r \in \mathbb{R}^+$, η τετραγωνική μορφή μπορεί να υπολογιστεί μέσω διαδοχικών γινομένων πίνακα με διάνυσμα. Για οποιαδήποτε άλλη συνάρτηση, για πίνακες μικρής διάστασης, η ποσότητα $x^T f(A) x$ μπορεί να υπολογιστεί απ' ευθείας χρησιμοποιώντας κλασικές υπολογιστικές μεθόδους. Τέτοιοι αλγόριθμοι που στηρίζονται σε αυτές τις τεχνικές έχουν κόστος τουλάχιστον $\mathcal{O}(p^3)$ πράξεων, για πυκνούς πίνακες. Στις εφαρμογές, καθώς η διάσταση p των απαιτούμενων πινάκων αυξάνει, είναι πολύ σημαντική η ανάπτυξη αποτελεσματικών μεθόδων που μειώνουν την τάξη της απαιτούμενης πολυπλοκότητας. Οι εκτιμήσεις των τετραγωνικών μορφών $x^T f(A) x$ που θα χρησιμοποιήσουμε για τον υπολογισμό των εκτιμητών του ίχνους στηρίζονται στη μέθοδο της παρεκβολής (extrapolation) των ροπών του πίνακα A , όπως παρουσιάζεται στην εργασία [28]. Η μέθοδος της παρεκβολής των ροπών ενός πίνακα A έχει εκτενώς μελετηθεί στις εργασίες [15], [17], [16], [19], [20], [29]. Φράγματα για τις ροπές ενός πίνακα A δίνονται στην εργασία [18].

Άλλες μέθοδοι που έχουν αναπτυχθεί στην βιβλιογραφία παρουσιάζουν διαδικασίες εκτίμησης του $\text{Tr}(f(A))$ για συμμετρικούς, θετικά ορισμένους πίνακες χρησιμοποιώντας κανόνες αριθμητικής ολοκλήρωσης (τύποι Gauss, Gauss Lobatto, Gauss Radau) (βλ. [5], [32]), modified moments (βλ. [32], [46]), και τεχνικές Monte Carlo (βλ. [3]). Η κύρια ιδέα αυτών των μεθόδων είναι να μετατρέψουν την τετραγωνική μορφή $x^T f(A)x$ σε ένα Riemann-Stieltjes ολοκλήρωμα και στη συνέχεια να χρησιμοποιήσουν κανόνες ολοκλήρωσης για να προσεγγίσουν το ολοκλήρωμα, μέσω της θεωρίας των ορθογωνίων πολυωνύμων και της μεθόδου Lanczos. Η απαιτούμενη πολυπλοκότητα αυτών των μεθόδων είναι $\mathcal{O}(jp^2)$ για πυκνούς πίνακες, όπου j είναι το πλήθος των επαναλήψεων της μεθόδου Lanczos. Για $f(A) = A^{-1}$, μέθοδοι προσέγγισης του $x^T A^{-1}x$ δίνονται επίσης στην εργασία [58], για κάθε μη ιδιάζοντα μιγαδικό πίνακα A και μιγαδικό διάνυσμα x . Αυτές οι μέθοδοι βασίζονται στην εκτίμηση της ποσότητας $x^T y$, όπου το y είναι η λύση του γραμμικού συστήματος $Ay = x$, μέσω μίας αποτελεσματικής προσέγγισης, χρησιμοποιώντας επαναλήψεις της μεθόδου biconjugate gradient. Για $f(A) = A^{-1}$, μία άλλη προσέγγιση αναπτύχθηκε επίσης στην εργασία [64].

Η εκτίμηση του ίχνους συνάρτησης πίνακα έχει πολλές εφαρμογές γι αυτό και έχει μελετηθεί εκτενώς (βλ. [32] και οι βιβλιογραφικές αναφορές που παρατίθενται εκεί). Παρακάτω, θα δούμε κάποιες εφαρμογές του υπολογισμού του ίχνους $\text{Tr}(f(A))$ σε περιοχές των μαθηματικών, της στατιστικής και της φυσικής (βλ. επίσης [55]).

– Εφαρμογές του $\text{Tr}(f(A))$

Η εκτίμηση του ίχνους συναρτήσεων πινάκων εμφανίζεται σε τομείς όπως οι εξής.

- Φυσική-χβαντική χημεία: για την εκτίμηση της πυκνότητας density of states (βλ. [53]).
- Ανάλυση δικτύων: Για έναν πίνακα γειτνίασης A , ο “Estrada index”

ορίζεται ως το ίχνος $\text{Tr}(\exp(A))$ (βλ. [24], [26]).

- Ορίζουσα του πίνακα A : Ο υπολογισμός της ορίζουσας ενός πίνακα A χρειάζεται στην lattice κβαντική χρωμοδυναμική και σε άλλες περιοχές. Εκτιμήσεις για την ορίζουσα ενός πίνακα ($\det(A)$) δίνονται εκτιμώντας το ίχνος $\text{Tr}(\log(A))$ (βλ. [32], [41]).
- Μερικό άθροισμα ιδιοτιμών: Στη φυσική στερεάς κατάστασης (solid state physics) ο υπολογισμός της συνολικής ενέργειας της ηλεκτρονικής δομής απαιτεί τον υπολογισμό των αθροισμάτων μερικών ιδιοτιμών. Για κατάλληλη επιλογή της συνάρτησης f , η εκτίμηση του $\text{Tr}(f(A))$ οδηγεί στην εκτίμηση του μερικού αθροίσματος των ιδιοτιμών ενός πίνακα A (βλ. [4]).
- Μηχανική μάθηση: Το πρόβλημα βελτιστοποίησης που προκύπτει κατά την εκπαίδευση μίας Gaussian διαδικασίας απαιτεί τον υπολογισμό του $\text{Tr}(T^{-1}S)$, για κατάλληλους πίνακες T, S (βλ. [12]).

Πιο συγκεκριμένα, το ίχνος $\text{Tr}(A^r)$, $r \in \mathbb{R}$, έχει λάβει ιδιαίτερη προσοχή, λόγω των πολλών εφαρμογών στις οποίες εμφανίζεται. Παρακάτω, αναφέρονται συγκεκριμένες εφαρμογές, με ιδιαίτερη έμφαση στο $\text{Tr}(A^{-1})$.

– **Εφαρμογές του $\text{Tr}(A^r)$, $r \in \mathbb{R}$**

Η εκτίμηση του ίχνους δυνάμεων πινάκων χρειάζεται σε διάφορους τομείς των μαθηματικών. Πιο συγκεκριμένα

- Ανάλυση Δικτύων: καταμέτρηση τριγώνων σε ένα γράφημα

Κατά την ανάλυση ένα πολύπλοκου δίκτυου, ένα σημαντικό πρόβλημα είναι να υπολογιστεί ο συνολικός αριθμός των τριγώνων (δηλαδή των συνδεδεμένων κορυφών ανά τριάδες) ενός συνδεδεμένου απλού γραφήματος. Αυτός ο αριθμός είναι ίσος με το $\text{Tr}(A^3)/6$, όπου A είναι ο

πίνακας γειτνίασης (adjacency matrix) του γράφου [1]. Για πολλά δίκτυα, ακόμη και αν ο πίνακας A είναι αραιός, ο πίνακας A^3 μπορεί να είναι αρκετά πυκνός και, ως εκ τούτου, δεν είναι δυνατόν να υπολογιστεί άμεσα αυτό το ίχνος.

Μέσω γραφημάτων επιτυγχάνεται η μοντελοποίηση πολλών φαινομένων (πχ WorldWideWeb, κοινωνικά δίκτυα, δίκτυα αλληλεπίδρασης πρωτεϊνών). Η καταμέτρηση τριγώνων στο γράφημα έχει πολλές εφαρμογές όπως για την ανίχνευση της ανεπιθύμητης αλληλογραφίας (spam detection). Επίσης, στο πλαίσιο της κοινωνικής δικτύωσης, καθώς οι φίλοι των φίλων τείνουν να γίνονται και φίλοι μεταξύ τους, ο αριθμός αυτός μπορεί να χρησιμοποιηθεί στη σύσταση συνδέσμου ή για πρόταση φιλίας στα μέσα κοινωνικής δικτύωσης. Η εκτίμηση του συνολικού αριθμού των τριγώνων σε ένα γράφημα είναι πολύ σημαντική καθώς η καταμέτρηση τριγώνων σε μεγάλα γραφήματα που συνήθως συναντούνται σε πραγματικές εφαρμογές (πχ WorldWideWeb, Facebook, Email) είναι υπολογιστικά ακριβή.

- Θεωρία Αριθμών και Συνδυαστική: Euler congruence.

Το ίχνος δυνάμεων πινάκων συνδέεται με την Euler congruence [65], ένα σημαντικό φαινόμενο στα μαθηματικά, που δηλώνει ότι $\text{Tr}(A^{p^r}) \equiv \text{Tr}(A^{p^{r-1}}) \pmod{p^r}$, για κάθε πίνακα A , κάθε πρώτο αριθμό p , και για κάθε $r \in \mathbb{N}$.

Η ποικιλομορφία των αποδείξεων που υπάρχουν για την Euler congruence δείχνει την καθολικότητά της και το ρόλο της σε διάφορους κλάδους των μαθηματικών.

- Στατιστική: Προσδιορισμός του κλασικού κριτηρίου βελτιστοποίησης (specification of classical optimality criteria).

Στον βέλτιστο σχεδιασμό πειραμάτων (βλ. [51]) ο τελικός στόχος κάθε βέλτιστου κριτηρίου είναι να μετρήσει την επίδραση ενός θετικά ημιο-

ρισμένου πίνακα C διάστασης s . Ένα από τα πιο σημαντικά κριτήρια είναι το κριτήριο της μέσης διασποράς $\Phi_{-1}(C) = (\text{Tr}(C^{-1})/s)^{-1}$, αν ο πίνακας C είναι θετικά ορισμένος.

Κριτήρια που είναι αναλλοίωτα κάτω από αναπαραμετροποίηση χάνουν το ενδιαφέρον τους εάν οι παράμετροι που μας ενδιαφέρουν έχουν μια σαφή φυσική σημασία. Το παραπάνω κριτήριο μέσης διακύμανσης παρέχει μια εύλογη εναλλακτική λύση. Γενικότερα, για θετικά ορισμένους πίνακες C , ο πίνακας Φ_p μπορεί να οριστεί για κάθε πραγματική παράμετρο p από τη σχέση $\Phi_p(C) = (\text{Tr}(C^p)/s)^{1/p}$, για $p \neq 0, \pm\infty$.

- Δυναμικά συστήματα: προσδιορισμός των αναλλοίωτων (determination of their invariants).

Οι αναλλοίωτες των δυναμικών συστημάτων περιγράφονται χρησιμοποιώντας το ίχνος δυνάμεων πινάκων (για παράδειγμα στη μελέτη τους αριθμούς Lefschetz) [65].

- p -αδική ανάλυση: καθορισμός του διανύσματος Witt.

Για οποιονδήποτε πίνακα M και κάθε πρώτο αριθμό p οι είσοδοι του μοναδικού διανύσματος Witt που αποτελείται από p -αδικούς ακέρατους αριθμούς εκφράζονται από τα ίχνη δυνάμεων πινάκων M [65].

- Θεωρία Πινάκων: ακραίες ιδιοτιμές.

Υπάρχουν πολλές εφαρμογές στη θεωρία πινάκων και στη αριθμητική γραμμική άλγεβρα. Για παράδειγμα, για την προσέγγιση της ελάχιστης και της μέγιστης ιδιοτιμής ενός συμμετρικού πίνακα A , μέσω μιας διαδικασίας που βασίζεται στην εκτίμηση του ίχνου των πινάκων A^n και A^{-n} , $n \in \mathbb{N}$, που προτείνεται στην εργασία [49].

- Διαφορικές εξισώσεις: επίλυση της εξίσωσης πίνακα Lyapunov.

Αυτές οι εξισώσεις μπορούν να λυθούν με τη χρήση πολυωνύμων πινάκων και των χαρακτηριστικών πολυωνύμων όπου απαιτείται ο υπολογι-

σμός του ίχνους δυνάμεων πινάκων [22].

– Εφαρμογές του $\text{Tr}(A^{-1})$

Εκτιμήσεις του $\text{Tr}(A^{-1})$ [7] έχουν εφαρμογές σε τομείς όπως οι παρακάτω.

- Στατιστική: προβλήματα σχεδιασμού στάθμισης (weighing design problems.)

Στους σχεδιασμούς στάθμισης, ο πίνακας σχεδιασμού \tilde{X} , καθορίζεται ελαχιστοποιώντας το ίχνος του αντιστρόφου του συμμετρικού και θετικά ορισμένου πίνακα $\tilde{X}^T \tilde{X}$ (A -optimality criterion). Μέθοδοι για την κατασκευή τέτοιων σχεδιασμών στάθμισης βέλτιστης χημικής ισορροπίας περιγράφονται, για παράδειγμα, στις εργασίες [62], [21].

- Στατιστική: Προσδιορισμός του κλασικού κριτηρίου βελτιστοποίησης. (specification of classical optimality criteria.)

Στον βέλτιστο σχεδιασμό πειραμάτων (βλ. [51]) ο τελικός στόχος κάθε βέλτιστου κριτηρίου είναι να μετρήσει την επίδραση ενός θετικά ημιορισμένου πίνακα C διάστασης s . Ένα από τα πιο σημαντικά κριτήρια είναι το κριτήριο της μέσης διασποράς $\Phi_{-1}(C) = (\text{Tr}(C^{-1})/s)^{-1}$, αν ο πίνακας C είναι θετικά ορισμένος.

Κριτήρια που είναι αναλλοίωτα κάτω από αναπαραμετροποίηση χάνουν το ενδιαφέρον τους εάν οι παράμετροι που μας ενδιαφέρουν έχουν μια σαφή φυσική σημασία. Το παραπάνω μέσο κριτήριο διακύμανσης παρέχει μια εύλογη εναλλακτική λύση. Γενικότερα, για θετικά ορισμένους πίνακες C , ο πίνακας Φ_p μπορεί να οριστεί για κάθε πραγματική παράμετρο p από τη σχέση $\Phi_p(C) = (\text{Tr}(C^p)/s)^{1/p}$, για $p \neq 0, \pm\infty$.

- Φυσική: Υπολογισμός των quark βρόχων στο πλέγμα κβαντικής χρωμοδυναμικής (QCD).

Πολλά φυσικά συστήματα, κλασικά ή μηχανικά, απαιτούν τον υπολογισμό του ίχνους ενός πίνακα M^{-1} , όταν η διάσταση του M αυξάνεται

γρήγορα με τις φυσικές μεταβλητές του προβλήματος (βλ. [23]). Ο πίνακας quark στο πλέγμα υπολογισμών της κβαντικής χρωμοδυναμικής (QCD) εμπίπτει σε αυτή την κατηγορία. Σε αυτές τις περιπτώσεις, απαιτείται ιδιαίτερα η εκτίμηση του ίχνους των πινάκων των οποίων οι αντίστροφοι κυριαρχούνται από τη διαγώνιο και τα στοιχεία που είναι κοντά στη διαγώνιο.

- Κρύσταλλοι: selection of measurement directions in elastic strain determination of single crystals.

Οι ποσότητες $\text{Tr}(P^{-1})$, όπου P ένας κατάλληλα κατασκευασμένος, συμμετρικός και θετικά ορισμένος πίνακας, είναι κατάλληλα εργαλεία για να αποφασιστεί εάν οι, όπως αποκαλούνται, measurement directions έχουν επιλεγεί σωστά (βλ. [48]).

- Ανάλυση δικτύων και η θεωρία γραφημάτων: Υπολογισμός της ποσότητας resolvent Estrada index η οποία ορίζεται ως το ίχνος $\text{Tr}((I_p - aA)^{-1})$ (βλ. [24], [10], [11]).
- Ανάλυση Κινδύνου (risk analysis): Η ποσοτικοποίηση της αβεβαιότητας (uncertainty quantification) στην ανάλυση κινδύνου απαιτεί το ίχνος του αντιστρόφου κατάλληλων πινάκων συνδιασποράς (inverse covariance matrices) για τον υπολογισμό του βαθμού εμπιστοσύνης στην ποιότητα των δεδομένων (βλ. [8], [9], [42], [59]).
- Για περισσότερες εφαρμογές σχετικά με το ίχνος $\text{Tr}(A^{-1})$, ή εν γένει $\text{Tr}(A^r)$, $r \in \mathbb{R}$, ο αναγνώστης μπορεί να ανατρέξει στις εργασίες [32], [3], [54], [57], [43].

Παρακάτω, θα παρουσιασθούν οι εκτιμητές του ίχνους συνάρτησης του πίνακα A .

Συγκεκριμένα, στο Κεφάλαιο 1 θα παρουσιαστεί το Θεώρημα του Hutchinson, ένα στοχαστικό αποτέλεσμα όπου βασίζονται οι στοχαστικοί εκτιμητές

του ίχνους, καθώς επίσης και σχετικές ανισότητες. Επίσης αναπτύσσονται οι στοχαστικοί εκτιμητές του ίχνους, οι οποίοι συγκρίνονται ως προς τις ιδιότητές τους. Τέλος, παρουσιάζονται διαστήματα εμπιστοσύνης.

Στο Κεφάλαιο 2 γίνεται μελέτη του σφάλματος των στοχαστικών εκτιμητών Hutchinson καθώς και ανάλυση στατιστικών σχεδιασμών με σκοπό την βελτίωση της ποιότητάς τους.

Στο Κεφάλαιο 3 παρουσιάζονται ποικίλα παραδείγματα και εφαρμογές των στατιστικών μεθόδων στους εκτιμητές για την προσέγγιση του ίχνους διαφορών συναρτήσεων πινάκων που απαιτούνται σε πραγματικές εφαρμογές.

Στον παρόν κεφάλαιο της εργασίας και σε όλα τα επόμενα, το $\|x\|$ συμβολίζει την Ευκλείδεια νόρμα ενός διανύσματος x , η $\|A\|$ είναι η φασματική νόρμα ενός πίνακα A , (\cdot, \cdot) συμβολίζει το εσωτερικό γινόμενο, ο εκθέτης T συμβολίζει τον ανάστροφο πίνακα και το σύμβολο \simeq σημαίνει “περίπου ίσο με” και το N αντιπροσωπεύει το πλήθος των διανυσμάτων στο δείγμα (sample vectors) ώστε να υπολογιστούν οι στοχαστικοί εκτιμητές. Όπως είδαμε, η μέγιστη και η ελάχιστη ιδιοτιμή ενός πίνακα συμβολίζονται με λ_1 και λ_p , αντίστοιχα.

Κεφάλαιο 1

Στοχαστικοί εκτιμητές του ίχνους

Παρουσιάζουμε στη συνέχεια ένα χρήσιμο Θεώρημα που έχει αναπτυχθεί στην εργασία [40], στο οποίο έχουν στηριχθεί οι υπάρχουσες μέθοδοι Monte Carlo για την εκτίμηση του ίχνους πινάκων (βλ. [3], [9], [16], [17], [32], [46]).

1.1 Το θεώρημα του Hutchinson

Για έναν πίνακα $A \in \mathbb{R}^{p \times p}$, η ακριβής τιμή του ίχνους του $f(A)$ του δίνεται από το άθροισμα

$$\text{Tr}(f(A)) = \sum_{i=1}^p e_i^T f(A) e_i,$$

όπου $e_i, i = 1, \dots, p$, είναι οι p γραμμικά ανεξάρτητες στήλες του $p \times p$ ταυτοτικού πίνακα. Η τελευταία σχέση μας οδηγεί ώστε να θεωρήσουμε στοχαστικό εκτιμητή του ίχνους του πίνακα $f(A)$, ο οποίος να δίνεται αντικαθιστώντας αυτά τα p γραμμικά ανεξάρτητα διανύσματα από ένα διάνυσμα με στοχαστικά ανεξάρτητες εισόδους.

Ένας στοχαστικός εκτιμητής του ίχνους $\text{Tr}(I - A)$ δόθηκε αρχικά από τον Girard (βλ. [30]). Αυτός ο εκτιμητής είναι αμερόληπτος, και έχει σχετι-

κό σφάλμα που φράσσεται από την ποσότητα $(2/p)^{1/2}$, επομένως αποτελεί “εφικτή” εναλλακτική του ακριβούς υπολογισμού του ίχνους $\text{Tr}(I - A)$.

Παρακάτω, θα δούμε έναν εκτιμητή του ίχνους του πίνακα A (βλ. [40]) που υπολογίζεται σχηματίζοντας το γινόμενο $x^T Ax$, όπου x είναι ένα διάνυσμα με στοιχεία p ανεξάρτητα δείγματα από μία διακριτή τυχαία μεταβλητή που παίρνει τις τιμές 1 και -1 με ίση πιθανότητα $1/2$.

Λήμμα 3.

Έστω $A \in \mathbb{R}^{p \times p}$ συμμετρικός πίνακας και έστω $x = (x_1, \dots, x_n)^T$ ένα διάνυσμα με εισόδους p ανεξάρτητα δείγματα ($E(x_i x_j) = E(x_i)E(x_j)$) από μία τυχαία μεταβλητή X με μέσο μηδέν ($E(x_i) = 0$) και διασπορά σ^2 ($E(x_i^2) = \sigma^2$). Τότε

$$E((x, Ax)) = \sigma^2 \text{Tr}(A),$$

και

$$\text{Var}((x, Ax)) = 2\sigma^4 \sum_{i \neq j} (A)_{ij}^2 + [E(X^4) - \sigma^4] \left(\sum_i (A)_{ii}^2 \right),$$

όπου $E(\cdot)$ και $\text{Var}(\cdot)$ είναι η μέση τιμή και η διασπορά αντίστοιχα.

Απόδειξη

Για την τετραγωνική μορφή $x^T Ax$ έχουμε

$$(x, Ax) = \sum_{i,j} a_{ij} x_i x_j.$$

Επομένως,

$$\begin{aligned} E((x, Ax)) &= E\left(\sum_{i,j} a_{ij} x_i x_j\right) = \sum_{i,j} a_{ij} E(x_i x_j) = \\ &= \sum_i a_{ii} E(x_i^2) + 2 \sum_{i < j} a_{ij} E(x_i x_j), \end{aligned}$$

καθώς ο A είναι συμμετρικός και επομένως $a_{ij} = a_{ji}$.

Όμως, εφόσον το διάνυσμα $x = (x_1, \dots, x_n)^T$ έχει στοιχεία ανεξάρτητα δείγματα από την τυχαία μεταβλητή X , δηλαδή τα $x_i, x_j, i, j = 1, \dots, p$ είναι ανεξάρτητα μεταξύ τους θα έχουμε $E(x_i x_j) = E(x_i)E(x_j)$.

Επιπλέον, αυτά τα δείγματα έχουν μέσο μηδέν, δηλαδή $E(x_i) = 0$ και διασπορά σ^2 , δηλαδή $E(x_i^2) = \sigma^2$.

Επομένως $E(x_i x_j) = E(x_i)E(x_j) = 0$.

Χρησιμοποιώντας τις δύο τελευταίες σχέσεις στο παραπάνω άθροισμα έχουμε

$$\begin{aligned} E((x, Ax)) &= \sum_i a_{ii} E(x_i^2) + 2 \sum_{i < j} a_{ij} E(x_i x_j) = \\ &= \sum_i a_{ii} \sigma^2 + 0 = \sigma^2 \sum_i a_{ii} = \sigma^2 \text{Tr}(A). \end{aligned}$$

Κατ' αρχήν ισχύει για την διασπορά ότι

$$\text{Var}((x, Ax)) = E((x, Ax)^2) - (E((x, Ax)))^2.$$

Θα έχουμε

$$E((x, Ax)^2) = E\left(\left(\sum_{i,j} a_{ij} x_i x_j\right)\left(\sum_{i,j} a_{ij} x_i x_j\right)\right) = E\left(\sum_{i,j} \sum_{r,s} a_{ij} a_{rs} x_i x_j x_r x_s\right).$$

Παρατηρούμε ότι κάθε όρος του αθροίσματος με το οποίο εκφράζεται το $c_1^2(x)$ έχει τη μορφή $C x_1^{n_1} \cdots x_k^{n_k}$, με $x_i \neq x_j$ για $i \neq j$, $n_1 + \cdots + n_k = 4$, και όπου C είναι ένας συντελεστής ο οποίος είναι γινόμενο 2 στοιχείων του πίνακα A . Τότε, το $E(c_1^2(x))$ υπολογίζεται λαμβάνοντας υπ' όψιν ότι, αφού τα x_i είναι ανεξάρτητες τυχαίες μεταβλητές, η αναμενόμενη τιμή ενός γινομένου αυτών ισούται με το γινόμενο των αναμενόμενων τιμών, και ότι, για κάθε i , $E(x_i^m)$ ισούται με σ^2 αν το m είναι άρτιος αριθμός, και με 0 αν το m είναι περιττός αριθμός.

Επομένως, θα έχουμε

$$\begin{aligned}
 E((x, Ax)^2) &= E\left(\sum_{i,j} \sum_{r,s} a_{ij} a_{rs} x_i x_j x_r x_s\right) = \\
 &= 2 \sum_{i \neq j} a_{ij}^2 E(x_i^2 x_j^2) + \sum_{i \neq j} a_{ii} a_{jj} E(x_i^2 x_j^2) + \sum_i a_{ii}^2 E(x_i^4) + 0 = \\
 &= 2\sigma^4 \sum_{i \neq j} a_{ij}^2 + \sigma^4 \sum_{i \neq j} a_{ii} a_{jj} + E(X^4) \sum_i a_{ii}^2.
 \end{aligned}$$

Επίσης, έχουμε ότι

$$(E((x, Ax)))^2 = (\sigma^2 \sum_i a_{ii})^2 = \sigma^4 \sum_{i,j} a_{ii} a_{jj}.$$

Από τις δύο τελευταίες σχέσεις προκύπτει για την διασπορά

$$\begin{aligned}
 Var((x, Ax)) &= E((x, Ax)^2) - (E((x, Ax)))^2 = \\
 &= 2\sigma^4 \sum_{i \neq j} a_{ij}^2 + \sigma^4 \sum_{i \neq j} a_{ii} a_{jj} + E(X^4) \sum_i a_{ii}^2 - \sigma^4 \sum_{i,j} a_{ii} a_{jj} = \\
 &= 2\sigma^4 \sum_{i \neq j} a_{ij}^2 - \sigma^4 \sum_i a_{ii}^2 + E(X^4) \sum_i a_{ii}^2 \\
 &= 2\sigma^4 \sum_{i \neq j} a_{ij}^2 + (E(X^4) - \sigma^4) \sum_i a_{ii}^2,
 \end{aligned}$$

$$\text{καθώς } \sigma^4 \sum_{i \neq j} a_{ii} a_{jj} - \sigma^4 \sum_{i,j} a_{ii} a_{jj} = -\sigma^4 \sum_i a_{ii}^2. \quad \blacksquare$$

Από την πρώτη σχέση στο Λήμμα 5 βλέπουμε ότι η τετραγωνική μορφή $x^T Ax$ είναι ένας αμερόληπτος εκτιμητής του ίχνους $\text{Tr}(A)$ αν $\sigma^2 = 1$ και από τη δεύτερη σχέση βλέπουμε ότι το $x^T Ax$ έχει την ελάχιστη διασπορά αν ελαχιστοποιείται η ποσότητα $E(X^4) - \sigma^4$. Όμως η ποσότητα $Var(X^2) = E(X^2 - \sigma^2)^2 = E(X^4) - \sigma^4$ είναι μη αρνητική, με ελάχιστη τιμή το μηδέν αν η μέση τιμή του X^2 είναι σ^2 . Αυτό μας οδηγεί στο επόμενο Θεώρημα.

Θεώρημα 2.

Έστω $A = (a_{ij}) \in \mathbb{R}^{p \times p}$ συμμετρικός πίνακας με $\text{Tr}(A) \neq 0$. Έστω X μία διακριτή τυχαία μεταβλητή με τιμές 1 και -1 με ίση πιθανότητα 0.5, και έστω x ένα τυχαίο διάνυσμα με εισόδους p ανεξάρτητα δείγματα από τη X (στο εξής θα λέμε $x \in X^p$). Τότε η ποσότητα (x, Ax) είναι αμερόληπτος εκτιμητής του ίχνους του πίνακα A , δηλαδή,

$$E((x, Ax)) = \text{Tr}(A),$$

και

$$\text{Var}((x, Ax)) = 2 \sum_{i \neq j} a_{ij}^2,$$

όπου $E(\cdot)$ και $\text{Var}(\cdot)$ είναι η μέση τιμή και η διασπορά αντίστοιχα.

Επιπλέον, η X είναι η μοναδική τυχαία μεταβλητή ανάμεσα στις τυχαίες μεταβλητές με μέσο μηδέν για την οποία η ποσότητα $x^T Ax$ είναι ένας αμερόληπτος εκτιμητής του ίχνους $\text{Tr}(A)$ με ελάχιστη διασπορά.

Παρατήρηση 1.

Είδαμε ότι το διάνυσμα x έχει στοιχεία ανεξάρτητες και ομοιόμορφα κατανοημένες (*i.i.d.*) τυχαίες μεταβλητές της κατανομής *Rademacher* (δηλαδή έχει εισόδους 1 ή -1 με ίση πιθανότητα $1/2$ - $\text{Pr}(x(i) = \pm 1) = 1/2$). Η κατανομή *Rademacher* προκύπτει από την κατανομή *Bernoulli* με τον γραμμικό μετασχηματισμό $x = 2y - 1$, όπου το x έχει κατανομή *Rademacher* και το y κατανομή *Bernoulli*.

Παρατήρηση 2. Εξετάζοντας τη διασπορά του εκτιμητή $\text{Var}((x, Ax))$ βλέπουμε ότι διαισθητικά μετράει πόση από την “ενέργεια” του πίνακα (δηλαδή την *Frobenius* νόρμα) είναι πάνω στη διαγώνιο. Βλέπουμε επίσης ότι για γενικό πίνακα, η μέθοδος του *Hutchinson* μπορεί να μην είναι αποτελεσματική καθώς η απόκλιση $\text{Var}((x, Ax))$ εξαρτάται από τα στοιχεία του πίνακα (εκτός από αυτά που βρίσκονται στην κύρια διαγώνιο). Επομένως, υπάρχουν πίνακες

για τους οποίους η απόκλιση είναι μεγάλη, και επομένως η μέθοδος καθίσταται αναποτελεσματική γι' αυτούς. Για παράδειγμα, η απόκλιση $Var((x, Ax))$ για τον πίνακα A με όλα τα στοιχεία του 1, ο οποίος είναι συμμετρικός θετικά ημιορισμένος, είναι $2(p^2 - p)$, ενώ το ίχνος $Tr(A)$ είναι μόνο p .

Αποδεικνύονται τα εξής άνω φράγματα για την διασπορά.

Πόρισμα 2.

Εστω ο πίνακας $A \in \mathbb{R}^{p \times p}$ συμμετρικός με $Tr(A) \neq 0$, η διακριτή τυχαία μεταβλητή X και το διάνυσμα x όπως στο Θεώρημα 2. Τότε

$$Var(x^T Ax) \leq 2[Tr(A^2) - \frac{(Tr(A))^2}{p}].$$

Απόδειξη

Παρατηρούμε ότι

$$\sum_i a_{ii}^2 - \frac{(Tr(A))^2}{p} = \sum_i (a_{ii} - \frac{Tr(A)}{p})^2 \geq 0$$

αφού

$$\begin{aligned} \sum_i (a_{ii} - \frac{Tr(A)}{p})^2 &= \sum_i (a_{ii}^2 + \frac{(Tr(A))^2}{p^2} - \frac{2a_{ii}Tr(A)}{p}) = \\ &= \sum_i a_{ii}^2 + p \frac{(Tr(A))^2}{p^2} - \frac{2Tr(A)}{p} \sum_i a_{ii} = \sum_i a_{ii}^2 - \frac{(Tr(A))^2}{p}. \end{aligned}$$

Επομένως,

$$Var(x^T Ax) = 2 \sum_{i \neq j} a_{ij}^2 = 2[\sum_{i,j} a_{ij}^2 - \sum_i a_{ii}^2] \leq 2[Tr(A^2) - \frac{(Tr(A))^2}{p}].$$

■

Πόρισμα 3.

Έστω η μεταβλητή X και το διάνυσμα x όπως στο Θεώρημα 2 και έστω ο $A \in \mathbb{R}^{p \times p}$ συμμετρικός και θετικά ημιορισμένος πίνακας με ιδιοτιμές που βρίσκονται μεταξύ 0 και 1. Τότε

$$\text{Var}(x^T Ax) \leq \frac{2}{p} \text{Tr}(A) \text{Tr}(I_p - A).$$

Απόδειξη

Θέτουμε $A = Q^T D Q$, όπου ο Q είναι ορθογώνιος πίνακας και ο D είναι διαγώνιος με εισόδους μεταξύ 0 και 1 (βλ. [39]). Έπεται ότι

$$\text{Tr}(A^2) = \text{Tr}(D^2) \leq \text{Tr}(D) = \text{Tr}(A).$$

Από το Πόρισμα 2 έχουμε

$$\begin{aligned} \text{Var}(x^T Ax) &\leq 2 \left[\text{Tr}(A^2) - \frac{(\text{Tr}(A))^2}{p} \right] \leq \\ &\leq 2 \left[\text{Tr}(A) - \frac{(\text{Tr}(A))^2}{p} \right] = \frac{2}{p} \text{Tr}(A) (p - \text{Tr}(A)) = \\ &= \frac{2}{p} \text{Tr}(A) (\text{Tr}(I_p) - \text{Tr}(A)) = \frac{2}{p} \text{Tr}(A) \text{Tr}(I_p - A), \end{aligned}$$

λόγω της γραμμικότητας της συνάρτησης του ίχνους. ■

Πόρισμα 4.

Έστω ο $A \in \mathbb{R}^{p \times p}$ συμμετρικός, θετικά ημιορισμένος πίνακας με ιδιοτιμές που βρίσκονται μεταξύ 0 και 1. Έστω η μεταβλητή X και το διάνυσμα x όπως στο Θεώρημα 2 και υποθέτουμε επιπλέον ότι $\text{Tr}(A) \leq \epsilon p$ για $\epsilon < 1$. Τότε

$$\frac{\text{Var}(x^T (I_p - A)x)}{(\text{Tr}(I_p - A))^2} \leq \frac{2\epsilon}{(1 - \epsilon)p}.$$

Συγκεκριμένα για $\epsilon = 0.5$ το σχετικό τυπικό σφάλμα της ποσότητας $x^T (I_p - A)x$ εκτιμώντας το $\text{Tr}(I_p - A)$ φράσσεται από την ποσότητα $(2/p)^{1/2}$.

Απόδειξη

Προκύπτει άμεσα από το Πόρισμα 3 αν θεωρήσουμε αντί για τον πίνακα A , τον πίνακα $I_p - A$. ■

1.2 Hutchinson εκτιμητής του ίχνους $\text{Tr}(f(A))$

Είδαμε στο εισαγωγικό κεφάλαιο πως αν ο A είναι συμμετρικός, τότε και ο πίνακας $f(A)$ είναι επίσης συμμετρικός (βλ. [35]).

Απο το Θεώρημα του Hutchinson, αφού ο $f(A)$ είναι συμμετρικός, έπεται ότι

$$\text{Tr}(f(A)) = E((x, f(A)x))$$

για $x \in X^p$, και

$$\text{Var}((x, f(A)x)) = 2 \sum_{i \neq j} (f(A))_{ij}^2.$$

Επομένως, εκτιμήσεις του $\text{Tr}(f(A))$ μπορούν να παραχθούν ακολουθώντας την παρακάτω διαδικασία. Αρχικά πραγματοποιούμε αρκετά πειράματα χρησιμοποιώντας εκτιμήσεις για την τετραγωνική μορφή $(x, f(A)x)$ μέσω της μεθόδου της παρεχβολής των ροπών που αναπτύσσεται στην εργασία [28] για ερμιτιανούς πίνακες A . Στο εξής, αυτές οι εκτιμήσεις θα συμβολίζονται με $e_f(x)$, δηλαδή

$$e_f(x) \simeq (x, f(A)x).$$

Στη συνέχεια, υπολογίζουμε την αναμενόμενη τιμή $E(e_f(x))$ του $e_f(x)$ για $x \in X^p$.

Στην πράξη, μπορούμε, λοιπόν, να έχουμε εκτιμητές του $\text{Tr}(f(A))$ υπολογίζοντας N εκτιμήσεις $e_f(x_i)$ της ροπής $(x_i^T f(A)x_i)$, για διαφορετικά διανύσματα $x_i \in X^p$ τα οποία αντλούνται από ένα σύνολο 2^p διανυσμάτων, και στη συνέχεια βρίσκοντας την μέση τιμή αυτών των N ποσοτήτων.

Έχουμε τον παρακάτω εκτιμητή του ίχνους.

Ορισμός 4.

Ο εκτιμητής του ίχνους *Hutchinson* για τον συμμετρικό πίνακα $f(A)$ ορίζεται ως

$$(1.1) \quad H_N = \frac{1}{N} \sum_{i=1}^N e_f(x_i),$$

όπου τα $x_i \in X^p$, είναι N ανεξάρτητα τυχαία διανύσματα με εισόδους *i.i.d.* *Rademacher* τυχαίες μεταβλητές.

Είναι φυσικό να περιμένουμε ότι με ένα κατάλληλο μέγεθος δείγματος N , η μέση τιμή των ποσοτήτων $e_f(x_i)$ είναι μία καλή εκτίμηση του ίχνους $\text{Tr}(f(A))$.

Παρομοίως, εκτιμήσεις για την διασπορά του εκτιμητή δίνονται από τη σχέση

$$(1.2) \quad v_f = \frac{\sum_{i=1}^N (e_f(x_i) - H_N)^2}{N - 1}.$$

Παρ' όλα αυτά, είδαμε ότι η μέθοδος του *Hutchinson* μπορεί να μην είναι αποτελεσματική σε ορισμένους πίνακες, λόγω της διασποράς του εκτιμητή που εξαρτάται από τα στοιχεία του πίνακα $f(A)$. Γι' αυτόν τον λόγο αποδείχθηκαν τα άνω φράγματα για την διασπορά που δίνονται στα Πορίσματα 2- 4. Στη συνέχεια, θα δούμε άλλους στοχαστικούς εκτιμητές, οι οποίοι θα συγκριθούν με τον *Hutchinson* εκτιμητή του ίχνους ως προς την διασπορά και ως προς άλλες ιδιότητες.

1.3 Άλλοι στοχαστικοί εκτιμητές

Εναλλακτικά, μπορούμε να υπολογίσουμε στοχαστικούς εκτιμητές θεωρώντας τα διανύσματα x_i από διαφορετικές κατανομές. Στις εργασίες [1], [2] παρουσιάζονται δύο επιπλέον στοχαστικοί εκτιμητές.

Πρόταση 1.

Έστω $A = (a_{ij}) \in \mathbb{R}^{p \times p}$ συμμετρικός πίνακας. Έστω x ένα τυχαίο διάνυσμα με εισόδους που είναι ανεξάρτητες και ομοιόμορφα κατανομημένες (*i.i.d.*) μεταβλητές της τυπικής κανονικής κατανομής. Τότε η ποσότητα (x, Ax) είναι αμερόληπτος εκτιμητής του ίχνους του πίνακα A , δηλαδή,

$$E((x, Ax)) = \text{Tr}(A),$$

και

$$\text{Var}((x, Ax)) = 2\|A\|_F^2,$$

όπου $\|\cdot\|_F$ είναι η *Frobenius* νόρμα.

Επομένως, έχουμε τον ακόλουθο εκτιμητή του ίχνους του πίνακα $f(A)$.

Ορισμός 5.

Ο *Gaussian* εκτιμητής του ίχνους για τον συμμετρικό πίνακα $f(A) \in \mathbb{R}^{p \times p}$ ορίζεται ως

$$(1.3) \quad G_N = \frac{1}{N} \sum_{i=1}^N e_f(x_i),$$

όπου τα x_i , είναι N ανεξάρτητα τυχαία διανύσματα με εισόδους που είναι *i.i.d.* μεταβλητές της τυπικής κανονικής κατανομής $N(0, 1)$.

Παρατηρούμε ότι η απόκλιση του *Gaussian* εκτιμητή είναι μεγαλύτερη από αυτήν για τον εκτιμητή *Hutchinson*. Όμως ο *Gaussian* εκτιμητής έχει το πλεονέκτημα ότι επιτυγχάνει καλύτερο άνω φράγμα του μεγέθους του δείγματος που απαιτείται για να εξασφαλιστεί η σύγκλιση [2]. Να σημειωθεί επίσης ότι υπολογιστικά είναι πιο ακριβή η κατασκευή δείγματος από την τυπική κανονική κατανομή σε σχέση με την κατανομή *Rademacher*.

Ο τρίτος εκτιμητής σχεδιάστηκε για να μειώνει την τυχειότητα στον αλγόριθμο κατασκευής όπως για παράδειγμα όσο αφορά την επιλογή των διανυσμάτων,

καθώς η τυχειότητα στον αλγόριθμο επηρεάζει την παράλληλη επεξεργασία του. Συγκρινόμενος με τους δύο προηγούμενους εκτιμητές, ο εκτιμητής Hutchinson έχει λιγότερη τυχειότητα από τον Gaussian εκτιμητή και απαιτεί μόνο προσθέσεις, αφαιρέσεις και όχι πολλαπλασιασμούς.

Πρόταση 2.

Έστω $A = (a_{ij}) \in \mathbb{R}^{p \times p}$ συμμετρικός πίνακας. Έστω j τυχαίος ακέραιος μεταξύ του 1 και του p , και έστω $x = \delta_j$ το j -οστό μοναδιαίο διάνυσμα (δηλαδή η j -οστή στήλη του ταυτοτικού πίνακα διάστασης p , I_p). Τότε η ποσότητα $p(x, Ax)$ είναι αμερόληπτος εκτιμητής του ίχνους του πίνακα A , δηλαδή,

$$E(p(x, Ax)) = \text{Tr}(A),$$

και

$$\text{Var}(p(x, Ax)) = 2 \sum_{i=1}^p A_{i,i}^2 - (\text{Tr}(A))^2, .$$

Επομένως, έχουμε τον ακόλουθο εκτιμητή του ίχνους του πίνακα $f(A)$.

Ορισμός 6.

Ο Μοναδιαίος διανυσματικός *Unit vector* εκτιμητής του ίχνους για τον συμμετρικό, θετικά ορισμένο πίνακα $f(A) \in \mathbb{R}^{p \times p}$ ορίζεται ως

$$(1.4) \quad U_N = \frac{p}{N} \sum_{i=1}^N e_f(x_i),$$

όπου τα x_i , είναι N ανεξάρτητα ομοιόμορφα τυχαία διανύσματα από τη βάση $\delta_1, \delta_2, \dots, \delta_p$.

Παρατήρηση 3. Οι τετραγωνικές μορφές $\delta_j^T f(A) \delta_j$ είναι απλά τα διαγώνια στοιχεία $(f(A))_{j,j}$, κι επομένως η αναμενόμενη τιμή του $\delta_j^T f(A) \delta_j$ είναι ακριβώς η ποσότητα $\text{Tr}(A)/p$. Σε αντίθεση με τις μεταβλητές από τις προηγούμενες κατανομές, αυτές οι τετραγωνικές μορφές δεν εξαρτώνται καθόλου

από τα στοιχεία του πίνακα $f(A)$ που δεν είναι πάνω στην κύρια διαγώνιό του, παρά μόνο από τα διαγώνια στοιχεία του. Επομένως η σύγκλιση του εκτιμητή του ίχνους U_N δεν εξαρτάται από τα υπόλοιπα στοιχεία του πίνακα.

1.4 Υπολογιστική πολυπλοκότητα - συμπεράσματα

Εν κατακλείδι, όλοι οι στοχαστικοί εκτιμητές του $\text{Tr}(f(A))$ ακολουθούν το ίδιο μοτίβο. Κατ'αρχάς επιλέγεται τυχαία ένα διάνυσμα x από μία δεδομένη κατανομή και υπολογίζεται η ποσότητα $x^T f(A)x$. Αυτή η διαδικασία επαναλαμβάνεται N φορές χρησιμοποιώντας i.i.d. δείγματα, και κατόπιν υπολογίζεται ο μέσος όρος των N ποσοτήτων.

Κατά την εκτέλεση του αλγορίθμου οι εκτιμήσεις e_f των τετραγωνικών μορφών $x^T f(A)x$ υπολογίζονται μέσω γινομένων πίνακα με διάνυσμα και διάνυσμα με διάνυσμα ώστε να αποφεύγεται ο απευθείας υπολογισμός του πίνακα $f(A)$, κάτι που θα ανέβαζε σημαντικά την πολυπλοκότητα σε τουλάχιστον κυβική.

Επομένως οι εκτιμητές του ίχνους απαιτούν τον υπολογισμό μερικών εσωτερικών γινομένων και γινομένων πίνακα με διάνυσμα. Αν N είναι το μέγεθος του δείγματος, τότε η πολυπλοκότητα για τον υπολογισμό των στοχαστικών εκτιμητών του $\text{Tr}(f(A))$ είναι της τάξης του $O(Np^2)$.

Συμπερασματικά, συγκρίνοντας τους στοχαστικούς εκτιμητές που παρουσιάστηκαν, παρατηρούμε τα εξής. Κατ'αρχάς η διασπορά του Gaussian εκτιμητή είναι μεγαλύτερη από τη διασπορά του Hutchinson εκτιμητή. Γενικά, ο Hutchinson εκτιμητής έχει την μικρότερη διασπορά σε σχέση με τους στοχαστικούς εκτιμητές που παρουσιάστηκαν. Όμως, ο Gaussian εκτιμητής έχει το πλεονέκτημα ότι έχει καλύτερο άνω φράγμα για το απαιτούμενο μέγεθος δείγματος N που απαιτείται για να εξασφαλιστεί η σύγκλιση, δηλαδή μπορούμε να

έχουμε σύγκλιση χρησιμοποιώντας μικρότερο δείγμα απ' ότι θα χρειαζόταν ο Hutchinson ή ο Unit vector εκτιμητής. Να σημειωθεί ότι υπολογιστικά είναι πιο ακριβή η κατασκευή δείγματος από την τυπική κανονική κατανομή $N(0, 1)$ σε σχέση με την κατανομή Rademacher. Ο Unit vector εκτιμητής από την άλλη μεριά σχεδιάστηκε για να μειώνει την τυχειότητα στον αλγόριθμο κατασκευής, δηλαδή όσον αφορά την επιλογή των διανυσμάτων, καθώς οι δύο πρώτοι εκτιμητές χρησιμοποιούν πολύ μεγάλους δειγματικούς χώρους με τον Hutchinson εκτιμητή να έχει λιγότερη τυχειότητα από τον Gaussian. Συγκεκριμένα, ο Gaussian εκτιμητής χρησιμοποιεί συνεχείς τυχαίες μεταβλητές από την τυπική κανονική κατανομή, επομένως αντλεί τα διανύσματα από ένα σύνολο με θεωρητικά άπειρο πλήθος διανυσμάτων. Ο Hutchinson εκτιμητής αντλεί τα διανύσματα από ένα σύνολο με 2^p διανύσματα, εφόσον αυτά είναι μήκους p (p θέσεις) που έχουν τιμή 1 ή -1 . Τέλος, ο Unit vector εκτιμητής αντλεί τα διανύσματά του από τις στήλες του πίνακα A , δηλαδή από ένα σύνολο με p διανυσμάτων. Επίσης, αξίζει να σημειωθεί ότι ο Hutchinson εκτιμητής απαιτεί μόνο προσθέσεις, αφαιρέσεις και όχι πολλαπλασιασμούς, ενώ η σύγκλιση του Unit vector εκτιμητή εξαρτάται μόνο από τα διαγώνια στοιχεία του πίνακα A .

1.5 Διαστήματα Εμπιστοσύνης

Η παρακάτω Πρόταση μας δίνει το διάστημα εμπιστοσύνης για τους εκτιμητές T_f του ίχνους του συνάρτησης συμμετρικού πίνακα (όπου T_f μπορεί να είναι ένας οποιοσδήποτε εκτιμητής από τους H_N, G_N).

Το ποσοστό της μαρτυρίας που απαιτείται για να δεχθούμε ότι ένα γεγονός είναι απίθανο να συμβεί κατά τύχη λέγεται *επίπεδο σημαντικότητας*. Όσο πιο χαμηλό είναι το επίπεδο σημαντικότητας, τόσο πιο ισχυρή είναι η μαρτυρία. Η επιλογή του επιπέδου σημαντικότητας είναι αυθαίρετη (τυχαία).

Έστω $Z_{a/2}$ να είναι το $a/2$ ποσοστιαίο σημείο της κανονικής κατανομής

$\mathcal{N}(0, 1)$, δηλαδή η τιμή της τυπικής κανονικής κατανομής επάνω από την οποία η πιθανότητα να έχουμε κάποια παρατήρηση είναι $a/2$.

Το ακόλουθο αποτέλεσμα είναι το κλασσικό διάστημα εμπιστοσύνης σχετικό με την πιθανότητα να έχουμε καλό εκτιμητή (βλ. [16]).

Πρόταση 3.

$$Pr\left(\left|\frac{T_f - \text{Tr}(f(A))}{\sqrt{\text{Var}((x, f(A)x))/N}}\right| < Z_{a/2}\right) = 1 - a,$$

όπου N είναι το πλήθος των δοκιμών, a είναι το επίπεδο σημαντικότητας, και $Z_{a/2}$ είναι η *critical value* της τυπικής κανονικής κατανομής που ορίστηκε παραπάνω.

Για επίπεδο σημαντικότητας $a = 0.01$, έχουμε $Z_{a/2} = 2.58$, και η Πρόταση 3 μας δίνει διάστημα εμπιστοσύνης για το $\text{Tr}(f(A))$ με πιθανότητα $100(1 - a)\% = 99\%$.

Έχουμε,

$$T_f - \frac{\sqrt{\text{Var}((x, f(A)x))}}{\sqrt{N}} Z_{a/2} < \text{Tr}(f(A)) < T_f + \frac{\sqrt{\text{Var}((x, f(A)x))}}{\sqrt{N}} Z_{a/2}.$$

Επίσης, περιμένουμε, για κάθε μέγεθος δείγματος, ο εκτιμητής του ίχνους T_f να βρίσκεται στο διάστημα

$$\text{Tr}(f(A)) - \frac{\sqrt{\text{Var}((x, f(A)x))}}{\sqrt{N}} Z_{a/2} < T_f < \text{Tr}(f(A)) + \frac{\sqrt{\text{Var}((x, f(A)x))}}{\sqrt{N}} Z_{a/2}$$

με πιθανότητα 99%.

Εάν, στην Πρόταση 3, το $\text{Var}((x, f(A)x))$ αντικατασταθεί με το v_f που δίνεται από τη σχέση (1.2), τότε παίρνουμε μία προσέγγιση του διαστήματος εμπιστοσύνης.

Για να υπολογίσουμε τον εκτιμητή Hutchinson, μπορούμε να μην θεωρήσουμε τυχαία διανύσματα με εισόδους i.i.d Rademacher μεταβλητές, αλλά τις στήλες ενός πίνακα Hadamard [9], [27]. Αυτά τα διανύσματα ικανοποιούν το Θεώρημα του Hutchinson και επιπλέον είναι μεταξύ τους ορθογώνια.

Ένας $p \times p$ πίνακας H είναι πίνακας Hadamard εαν τα στοιχεία του είναι $+1$ ή -1 και $H^T H = H H^T = pI_p$, όπου I_p είναι ο $p \times p$ ταυτοτικός πίνακας. Οι στήλες και οι γραμμές του πίνακα H είναι ανα δύο ορθογώνιες.

Οι ιδιότητες αυτού του πίνακα, με κύρια την ορθογωνιότητα, βοηθούν στην μείωση του σχετικού σφάλματος εκτίμησης και συνεπώς στην βελτίωση του εκτιμητή, όπως θα δούμε αναλυτικά στο επόμενο Κεφάλαιο.

Κεφάλαιο 2

Στοχαστικοί εκτιμητές του ίχνους μέσω πινάκων Hadamard

Στο προηγούμενο Κεφάλαιο αναπτύχθηκαν εκτιμητές του ίχνους συνάρτησης του πίνακα A υπολογίζοντας τη μέση τιμή των ποσοτήτων $x_i^T f(A)x_i$, όπου x_i είναι κατάλληλα διανύσματα (sample vectors) που σχηματίζουν τις στήλες ενός πίνακα X (sample matrix). Στον παρόν Κεφάλαιο (βλ. [27]), η επιλογή των διανυσμάτων x_i από τις στήλες ενός πίνακα Hadamard, καθώς επίσης και η χρήση πινάκων Hadamard για την κατασκευή του sample matrix X μελετώνται και εξετάζονται μέσω διάφορων παραδειγμάτων. Επιπλέον, εισάγεται ο συνδυασμός των πινάκων Hadamard με έναν στατιστικό σχεδιασμό όπως το central-composite design (CCD) για την κατασκευή του sample matrix X .

Έστω $A \in \mathbb{R}^{p \times p}$ ένας συμμετρικός πίνακας και f μία συνεχώς διαφορίσιμη συνάρτηση που ορίζεται στο φάσμα του πίνακα A . Είδαμε ότι σε πολλές εφαρμογές απαιτείται η γνώση του ίχνους του πίνακα $f(A)$.

Η εκτίμηση του ίχνους μίας συνάρτησης πίνακα χρειάζεται σε πολλούς τομείς

όπως η Φυσική, όπου εμφανίζεται στην κβαντική χημεία για την εκτίμηση για την εκτίμηση της πυκνότητας density of states [53]. Επίσης, στην ανάλυση δικτύων, ο δείκτης Estrada ορίζεται ως το ίχνος $\text{Tr}(\exp(A))$, για πίνακα γειτνίασης A (βλ. [24], [25]).

Η εκτίμησης επίσης του ίχνους $\text{Tr}(A^{-1})$ έχει μελετηθεί ιδιαίτερα στην βιβλιογραφία (βλ. [32], [5], [17] και τις αναφορές που παρατίθενται εκεί). Το ίχνος του αντιστρόφου ενός συμμετρικού πίνακα απαιτείται σε πολλές εφαρμογές σε διάφορους τομείς όπως η στατιστική, τα fractals, το πλέγμα κβαντικής χρωμοδυναμικής (lattice quantum chromodynamics), οι κρύσταλλοι, η ανάλυση δικτύων και η θεωρία γραφημάτων. Επίσης, στην ανάλυση δικτύων, ο δείκτης resolvent Estrada ορίζεται ως το ίχνος $\text{Tr}((I_p - aA)^{-1})$, όπου το a είναι μία κατάλληλη παράμετρος, όπου επιλέγεται συνήθως $a = 0.85/\lambda_1$ (βλ. [24]) και I_p ο ταυτοτικός πίνακας διάστασης p . Τέλος, στη μηχανική μάθησης, το πρόβλημα βελτιστοποίησης που προκύπτει κατά την εκπαίδευση μιας Gaussian διαδικασίας απαιτεί τον υπολογισμό του ίχνους $\text{Tr}(T^{-1}S)$, για κατάλληλους πίνακες T, S (βλ. [12]).

Επιπλέον, το ίχνος $\text{Tr}(f(A))$ για κατάλληλη επιλογή της συνάρτησης f , μπορεί να μας δώσει εκτιμήσεις για άλλες χρήσιμες ποσότητες όπως το μερικό άθροισμα των ιδιοτιμών ενός πίνακα A , που εμφανίζεται στην φυσική στερεάς κατάστασης (solid state physics) για τον υπολογισμό της ολικής ενέργειας μιας ηλεκτρονικής δομής (βλ. [4].)

Επίσης, η εκτίμηση του ίχνους $\text{Tr}(f(A))$ μπορεί να οδηγήσει στην εκτίμηση της ορίζουσας ενός πίνακα A , που συμβολίζεται ως $\det(A)$, καθώς ισχύει ότι $\text{Tr}(\log(A)) = \log(\det(A))$ για έναν συμμετρικό θετικά ορισμένο πίνακα A (βλ. [32]). Ως συνέπεια, η εκτίμηση του ίχνους $\text{Tr}(\log(A))$ οδηγεί στην εκτίμηση της ορίζουσας $\det(A)$ που εμφανίζεται σε πολλές εφαρμογές όπως τα fractals, οι κρύσταλλοι, το πλέγμα κβαντικής χρωμοδυναμικής, η γενικευμένη cross-validation και οι εφαρμογές τους (βλ. [13].)

Έχουν προταθεί στοχαστικοί εκτιμητές για το $\text{Tr}(f(A))$, που βασίζονται στον υπολογισμό της ποσότητας $\sum_{i=1}^N x_i^T f(A) x_i / N$ για διαφορετικά sample ve-

ctors x_i , όπου N είναι το μέγεθος του δείγματος. Η αναμενόμενη τιμή της παραπάνω ποσότητας ισούται με το $\text{Tr}(f(A))$.

Στη συνέχεια, επικεντρωνόμαστε στην κατάλληλη επιλογή του sample matrix $X = [x_1, \dots, x_N]$, όπου x_i , $i = 1, \dots, N$ είναι τα sample vectors. Οι στοχαστικοί εκτιμητές και η επιλογή στηλών από πίνακες Hadamard ως sample vectors συζητώνται στις Ενότητες 2.1 και 2.2. Στην Ενότητα 2.4 εισάγεται η χρησιμοποίηση των πινάκων Hadamard σε έναν central-composite σχεδιασμό για την κατασκευή του sample matrix X .

2.1 Εκτιμητές μέσω της μεθόδου του Hutchinson

Έστω $A \in \mathbb{R}^{p \times p}$ ένας συμμετρικός πίνακας. Όπως είδαμε στο εισαγωγικό κεφάλαιο, ο πίνακας $f(A)$ είναι επίσης συμμετρικός [35]. Εκτιμήσεις για το ίχνος του, $\text{Tr}(f(A))$, μπορούν να παραχθούν εφαρμόζοντας το αποτέλεσμα του Hutchinson που εισήχθη στην εργασία [40] (και παρουσιάζεται στο προηγούμενο κεφάλαιο και δίνεται και στην επόμενη Πρόταση), στο οποίο δίνεται ένας ελάχιστης διασποράς εκτιμητής του ίχνους ενός συμμετρικού πίνακα A .

Πρόταση 4.

Έστω $A \in \mathbb{R}^{p \times p}$ ένας συμμετρικός πίνακας με $\text{Tr}(f(A)) \neq 0$. Έστω $x \in \mathbb{R}^p$ ένα διάνυσμα του οποίου τα στοιχεία παίρνουν τις τιμές 1 και -1 με ίση πιθανότητα 0.5. Τότε το $(x, f(A)x)$ είναι ένας αμερόληπτος εκτιμητής για το $\text{Tr}(f(A))$. Ισχύει

$$E((x, f(A)x)) = \text{Tr}(f(A)) \quad \text{και} \quad \text{Var}((x, f(A)x)) = 2 \sum_{i \neq j} (f(A))_{ij}^2,$$

όπου $E(\cdot)$ και $\text{Var}(\cdot)$ συμβολίζουν την αναμενόμενη τιμή και τη διασπορά αντίστοιχα.

Επομένως, εκτιμητές για το $\text{Tr}(f(A))$ δίνονται υπολογίζοντας τη μέση τιμή των ποσοτήτων $(x_i, f(A)x_i)$ για N διανύσματα όπως περιγράφονται στην Πρόταση 4. Επομένως, για μέγεθος δείγματος N , έχουμε

$$(2.1) \quad \text{Tr}(f(A)) \simeq \frac{1}{N} \sum_{i=1}^N x_i^T f(A) x_i.$$

Για τον υπολογισμό του (2.1) τα ακόλουθα δύο προβλήματα πρέπει να αντιμετωπιστούν.

(i) Αποτελεσματικός υπολογισμός των τετραγωνικών μορφών $x_i^T f(A) x_i$.

Είναι πολύ σημαντικό να αποφεύγεται ο απευθείας υπολογισμός του πίνακα $f(A)$ στις τετραγωνικές μορφές $x_i^T f(A) x_i$. Μέθοδοι που προτείνονται για αυτό έχουν αναπτυχθεί στο [32] και βασίζονται σε κανόνες αριθμητικής ολοκλήρωσης Gauss (βλ. επίσης [33], [44], [46]). Μία επιπλέον μέθοδος που στηρίζεται στην παρεκβολή έχει αναπτυχθεί στις εργασίες [17],[29], [16] για την εκτίμηση της τετραγωνικής μορφής $x^T A^{-1} x$, και γενικότερα της $x^T A^q x$, $q \in \mathbb{Q}$. Χρησιμοποιώντας την ίδια μέθοδο, εκτιμήσεις για το $x^T f(A) x$ δίνονται στην εργασία [28]. Είτε η μέθοδος της παρεκβολής, είτε οι κανόνες ολοκλήρωσης Gauss μπορούν να επιλεγούν για την αποτελεσματική προσέγγιση των τετραγωνικών μορφών $x_i^T f(A) x_i$.

(ii) Κατάλληλη επιλογή του sample matrix X .

Η κατάλληλη επιλογή των στηλών του sample matrix X επηρεάζουν τις στοχαστικές εκτιμήσεις. Έως τώρα, στις περισσότερες μεθόδους, τυχαία διανύσματα x χρησιμοποιούνταν σαν δείγμα, με ανεξάρτητες και ομοιόμορφα κατανομημένες (i.i.d.) Rademacher τυχαίες μεταβλητές σαν εισόδους ($Pr(x(i) = \pm 1) = 1/2$) [36], όπως περιγράφεται στην Πρόταση 4. Κατασκευάζονται χρησιμοποιώντας την κατανομή Bernoulli με τον γραμμικό μετασχηματισμό $\chi = 2\psi - 1$, όπου το χ έχει την κατανομή Rademacher και το ψ έχει την κατανομή Bernoulli.

Στη συνέχεια, μελετάμε πως μία συγκεκριμένη επιλογή των sample vectors μπορεί να βελτιώσει την ποιότητα των εκτιμητών του $\text{Tr}(f(A))$. Ας θεωρήσουμε έναν sample matrix $X \in \mathbb{R}^{p \times N}$ του οποίου οι στήλες είναι τα sample

vectors x_i . Το παρακάτω Λήμμα προσδιορίζει το σφάλμα στην εκτίμηση του $\text{Tr}(f(A))$ [63].

Λήμμα 4.

Έστω $A \in \mathbb{R}^{p \times p}$ ένας συμμετρικός πίνακας και $X = [x_1, \dots, x_N] \in \mathbb{R}^{p \times N}$ ο *sample matrix*, όπου x_i , $i = 1, \dots, N$ όπως περιγράφονται στην Πρόταση 4.

$$(i) \text{Tr}(f(A)) \simeq \frac{1}{N} \sum_{i=1}^N x_i^T f(A) x_i = \text{Tr}(X^T f(A) X) / N.$$

(ii) Το απόλυτο σφάλμα της προσέγγισης που περιγράφεται στο (i) είναι $\text{Err} = |\text{Tr}(f(A)(I_p - S))|$, όπου $S = XX^T / N$.

Απόδειξη

(i) Έπεται άμεσα από την Πρόταση 4.

$$\begin{aligned} (ii) \text{Err} &= \left| \text{Tr}(f(A)) - \frac{\text{Tr}(X^T f(A) X)}{N} \right| = \\ &= \left| \text{Tr}(f(A)) - \frac{\text{Tr}(f(A) X X^T)}{N} \right| = |\text{Tr}(f(A)(I_p - S))|, \quad S = \frac{X X^T}{N}. \end{aligned}$$

■

Όσο ο πίνακας S πλησιάζει τον ταυτοτικό I_p , το σφάλμα μειώνεται. Καθώς το μέγεθος του δείγματος N πρέπει να είναι αρκετά μικρότερο από p ($N \ll p$), προκειμένου να έχουμε έναν αποτελεσματικό στοχαστικό εκτιμητή, η περίπτωση να συμπίπτει ο πίνακας S με τον ταυτοτικό πίνακα I_p δεν μπορεί να θεωρηθεί. Επομένως, ενδιαφερόμαστε για *sample vectors* x_i τέτοια ώστε $XX^T \simeq NI_p$. Στη συνέχεια, θα δούμε πώς μπορεί να κατασκευαστεί ο *sample matrix* X χρησιμοποιώντας πίνακες Hadamard κατά τέτοιον τρόπο ώστε $XX^T \simeq NI_p$.

Ορισμός 7.

Ένας πίνακας *Hadamard* $H \in \mathbb{R}^{p \times p}$ είναι ένας πίνακας με στοιχεία 1 ή -1 , που ικανοποιεί τη σχέση ορθογωνιότητας $H^T H = H H^T = pI_p$.

Από αυτόν τον ορισμό έπεται ότι ανά δύο διαφορετικές γραμμές ή στήλες ενός πίνακα Hadamard είναι ορθογώνιες, δηλαδή το εσωτερικό τους γινόμενο είναι μηδέν. Μπορεί να αποδειχθεί ότι αν ο H είναι ένας πίνακας Hadamard διάστασης p , τότε $p = 1, 2$ ή $p = 0 \pmod{4}$. Για περισσότερες πληροφορίες σχετικά με πίνακες Hadamard, ο αναγνώστης μπορεί να συμβουλευτεί το [37]. Ένας πίνακας Hadamard λέγεται ότι είναι κανονικοποιημένος, αν έχει στην πρώτη γραμμή και στην πρώτη στήλη του όλα μονάδες.

Μια ειδική κατηγορία Hadamard πινάκων, οι οποίοι λέγονται Sylvester Hadamard πίνακες, μπορούν εύκολα να κατασκευαστούν με τη χρήση του μοτίβου $H_{k+1} = \begin{pmatrix} H_k & H_k \\ H_k & -H_k \end{pmatrix}$ όπου H_k είναι πίνακας Hadamard διάστασης 2^k για κάθε μη αρνητικό ακέραιο αριθμό k . $H_0 = \begin{pmatrix} 1 \end{pmatrix}$ και $H_1 = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$ είναι οι τετριμμένες περιπτώσεις. Αυτή η κατασκευή είναι ισοδύναμη με το γινόμενο Kronecker $H_1 \otimes H_k$.

Στη συνέχεια, θα μελετήσουμε την περίπτωση του να θέσουμε ως στήλες του sample matrix X , στήλες (ή μέρος των στηλών) πινάκων Hadamard. Εάν θεωρήσουμε ως sample matrix X έναν πίνακα Hadamard διάστασης p ($X \equiv H$), τότε $XX^T = pI_p$ και επομένως η στοχαστική εκτίμηση θα γίνει ακριβής. Όμως, καθώς χρειαζόμαστε $N \ll p$, μελετάμε την αποτελεσματικότητα της χρήσης ορισμένων στηλών του πίνακα H .

2.2 Δειγματοληψία από Sylvester Hadamard πίνακες

Οι πίνακες Hadamard είχαν επίσης χρησιμοποιηθεί στην εργασία [63] για την εκτίμηση του $\text{Tr}(f(A))$ στην κατασκευή του sample matrix X , σε συνδυασμό με άλλους πίνακες (βλ. επίσης [62]). Επιπλέον, στην εργασία [1], οι πίνακες Hadamard προτάθηκαν για να περιστρέψουν τα ορθοκανονικά διανύσματα

σε έναν εκτιμητή του ίχνους, και, πιο πρόσφατα, τα οφέλη από τις στήλες πινάκων Hadamard συζητήθηκαν στην εργασία [56] για την εκτίμηση του ίχνους του αντιστρόφου πινάκων ειδικής δομής (special structured matrices). Τέλος, στην εργασία [9] οι στήλες των πινάκων Hadamard χρησιμοποιήθηκαν στον sample matrix X , αντί των διανυσμάτων Rademacher για την εκτίμηση της διαγωνίου ενός πίνακα. Σε αυτήν την εργασία μελετήθηκαν οι ιδιότητες του πίνακα XX^T και δόθηκε μια “εικόνα” σχετικά με τη χρήση δειγμάτων μεγέθους που είναι δύναμη του δύο. Στην εργασία [27], όπως αναπτύσσεται και στο παρόν Κεφάλαιο, αποδεικνύεται η χρήση μεγέθους δείγματος που είναι δύναμη του δύο.

Ας θεωρήσουμε ότι ο sample matrix X είναι ο $p \times N$ κύριος υποπίνακας ενός Sylvester Hadamard πίνακα, όπου $p > N = 2^i$, $i \in \mathbb{N}$. Το ακόλουθο Λήμμα δείχνει ότι ο πίνακας X μπορεί να διαχωριστεί σε blocks.

Λήμμα 5.

Έστω H_k ένας Sylvester Hadamard πίνακας διάστασης 2^k και έστω X ο $p \times 2^i$ κύριος υποπίνακας του, όπου $i \in \mathbb{N}$ τέτοιο ώστε $2^i < p$.

(i) Αν $p = 2^k$, τότε ο πίνακας X αποτελείται από 2^{k-i} blocks διάστασης $2^i \times 2^i$.

Κάθε block είναι ένας Sylvester Hadamard πίνακας H_i διάστασης 2^i .

(ii) αν $p \neq 2^k$, τότε $p \equiv r \pmod{2^i}$ για $r \in \mathbb{N}$, $0 \leq r < 2^i$. Σε αυτήν την περίπτωση, ο πίνακας X αποτελείται από $2^i \times 2^i$ Sylvester Hadamard πίνακες H_i επαυξημένοι με τις πρώτες r γραμμές του πίνακα H_i .

Απόδειξη

(i) Μπορεί να αποδειχθεί με επαγωγή και από την κατασκευή των Sylvester Hadamard πινάκων, καθώς $p = 2^k \equiv 0 \pmod{2^i}$ για οποιοδήποτε $i = 1, \dots, k$. Για $k = 1$ έχουμε τον 2×2 Sylvester Hadamard πίνακα $H_1 = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$ και κατασκευάζουμε τον $2^2 \times 2^2$ Sylvester Hadamard πίνακα

$$H_2 = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{pmatrix}. \text{ Στη συνέχεια, θα έχουμε τον } 2^3 \times 2^3 \text{ Sylve-}$$

$$\text{ster Hadamard πίνακα } H_3 = \left(\begin{array}{cccc|cccc} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 & 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 \\ \hline 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 \\ 1 & -1 & 1 & -1 & -1 & 1 & -1 & 1 \\ 1 & 1 & -1 & -1 & -1 & -1 & 1 & 1 \\ 1 & -1 & -1 & 1 & -1 & 1 & 1 & -1 \end{array} \right).$$

Σε αυτόν τον $2^3 \times 2^3$ πίνακα, κάθε $2^i \times 2^i$, $i = 1, 2$ block είναι ένας Sylvester Hadamard πίνακας H_i διάστασης 2^i . Ας υποθέσουμε ότι ισχύει για $k = n \in \mathbb{N}$. επομένως, θα έχουμε έναν $2^n \times 2^n$ πίνακα H_n του οποίου κάθε $2^i \times 2^i$ για $i = 1, 2, \dots, n$ block είναι ένας Sylvester Hadamard πίνακας H_i . Τότε, ο $2^{n+1} \times 2^{n+1}$ πίνακας $H_{n+1} = \begin{pmatrix} H_n & H_n \\ H_n & -H_n \end{pmatrix}$ μπορεί να κατασκευαστεί, ο οποίος περιέχει τους $2^i \times 2^i$ για $i = 1, 2, \dots, n+1$ Sylvester Hadamard πίνακες H_i ως blocks.

(ii) Καθώς $p \equiv r \pmod{2^i}$ για $r \in \mathbb{N}$, $0 \leq r < 2^i$, έχουμε ότι $p = t2^i + r$, για μία σταθερά $t \in \mathbb{N}$. Λαμβάνοντας υπόψιν την προαναφερθείσα ανάλυση για τον $t2^i \times 2^i$ κύριο υποπίνακα του πίνακα H_k , έπεται ότι ο $p \times 2^i$ κύριος υποπίνακας του πίνακα H_k αποτελείται από t Sylvester Hadamard πίνακες H_i επαυξημένους με τις πρώτες r γραμμές του πίνακα H_i . ■

Στη συνέχεια, στο Θεώρημα 3 δίνουμε μία απόδειξη σχετικά με την αποτελεσματικότητα της χρήσης μεγέθους δείγματος διάστασης που είναι δύναμη του δύο.

Θεώρημα 3.

Έστω $A \in \mathbb{R}^{p \times p}$ ένας συμμετρικός πίνακας. Αν θεωρήσουμε ως *sample matrix* X τον $p \times N$ κύριο υποπίνακα του Sylvester Hadamard πίνακα H_k διάστασης 2^k , για $N = 2^i$, $i < k$, τότε ο πίνακας $S = XX^T/N$ είναι αραιός με μονάδες στην κύρια διαγώνιο και σε κάθε $k2^i$ υπερ και υποδιαγώνιο, $\forall k \in \mathbb{N}$

τέτοιο ώστε $\kappa 2^i \leq p$. Ειδικότερα,

$$(i) \text{ αν } p = 2^k \text{ τότε ο πίνακας } \frac{XX^T}{N} \text{ είναι } \hat{I}_p = \begin{pmatrix} I_N & I_N & \cdots & I_N \\ I_N & I_N & \cdots & I_N \\ \vdots & \vdots & \ddots & \vdots \\ I_N & I_N & \cdots & I_N \end{pmatrix}.$$

(ii) αν $p \neq 2^k$, τότε υπάρχει $r \in \mathbb{N}$, $0 \leq r < 2^i$, τέτοιο ώστε $p \equiv r \pmod{2^i}$.

Τότε, ο πίνακας $\frac{XX^T}{N}$ είναι $\begin{bmatrix} \hat{I}_{p-r} & \hat{I}_{p-r}(:, 1:r) \\ \hat{I}_{p-r}(1:r, :) & I_r \end{bmatrix}$, όπου τα $\hat{I}_{p-r}(:, 1:r)$ και $\hat{I}_{p-r}(1:r, :)$ αντιπροσωπεύουν τους πίνακες με τις r πρώτες στήλες και τις r πρώτες γραμμές του πίνακα \hat{I}_{p-r} , αντίστοιχα.

Απόδειξη

(i) Ας θεωρήσουμε ότι $p = 2^k$. Αν $N = 2^i$, $i < k$, από το Λήμμα 5 έχουμε ότι κάθε $2^i \times 2^i$ block του $p \times 2^i$ πίνακα X είναι ένας Sylvester Hadamard πίνακας H_i . Επομένως, οι γραμμές του κάθε block είναι γραμμικά ανεξάρτητες δηλαδή το εσωτερικό τους γινόμενο είναι ίσο με μηδέν, και επομένως το αντίστοιχο στοιχείο του πίνακα XX^T στο κατάλληλο του $2^i \times 2^i$ block είναι μηδέν. Το εσωτερικό γινόμενο δύο ίδιων γραμμών είναι ίσο με N . Επομένως τα διαγώνια στοιχεία στο κατάλληλο $2^i \times 2^i$ block του πίνακα $S = \frac{XX^T}{N}$ είναι όλα μονάδες. Επομένως, ο πίνακας XX^T αποτελείται από $2^i \times 2^i$ blocks με διαγώνιες εισόδους όλα μονάδες, δηλαδή είναι της μορφής \hat{I}_p .

(ii) Στην περίπτωση που $p \neq 2^k$, τότε $p \equiv r \pmod{2^i}$, $r \in \mathbb{N}$, $0 \leq r < 2^i$. Επομένως, από το Λήμμα 5 έπεται άμεσα ότι ο πίνακας S είναι της μορφής \hat{I}_{p-r} επαυξημένος με τις πρώτες r στήλες του στα δεξιά και τις πρώτες r γραμμές του από κάτω, έχοντας τη μορφή $\begin{bmatrix} \hat{I}_{p-r} & \hat{I}_{p-r}(:, 1:r) \\ \hat{I}_{p-r}(1:r, :) & I_r \end{bmatrix}$. ■

Παρατηρούμε ότι, σαν αποτέλεσμα αυτής της επιλογής του N , “εξαφανίζονται” πολλά μη διαγώνια στοιχεία του S . Οι υπερ και υποδιαγώνιοι του πίνακα $S = \frac{XX^T}{N}$ επηρεάζουν το σφάλμα της εκτίμησης. Στην εργασία [9], επιση-

μαίνεται ότι όσο το πλήθος των στηλών ενός Sylvester Hadamard πίνακα αυξάνει, λιγότερες μη επιθυμητές διαγώνιοι εμφανίζονται στον πίνακα S .

Πόρισμα 5.

Έστω $A \in \mathbb{R}^{p \times p}$ ένας συμμετρικός πίνακας, f μία κατάλληλη λεία συνάρτηση που ορίζεται στο φάσμα του πίνακα A και $B = f(A)$. Αν θεωρήσουμε ως *sample matrix* X τον $p \times N$ κύριο υποπίνακα του Sylvester Hadamard πίνακα H_k , για $N = 2^i$, $i < k$, τότε το σφάλμα στην εκτίμηση του $\text{Tr}(B)$ επηρεάζεται μόνο από τα μη διαγώνια στοιχεία του πίνακα B .

Απόδειξη

(i) Ας θεωρήσουμε κατ' αρχάς την περίπτωση όπου $p = 2^k$. Έχουμε $\text{Tr}(B) \simeq \text{Tr}(X^T B X)/N$. Από το Λήμμα 5 ο *sample matrix* X αποτελείται από 2^{k-i} Sylvester Hadamard πίνακες διάστασης 2^i , δηλαδή $X = [H_i, H_i, \dots, H_i]^T$. Έστω B_{lj} , $l, j = 1, \dots, 2^{k-i}$ το lj τετραγωνικό block διάστασης N του πίνακα B . Έχουμε

$$\begin{aligned} \text{Tr}(X^T B X) &= \sum_{l,j=1}^{2^{k-i}} \text{Tr}(H_i^T B_{lj} H_i) = \sum_{l,j=1}^{2^{k-i}} \text{Tr}(B_{lj} H_i H_i^T) = \\ &= N \sum_{l,j=1}^{2^{k-i}} \text{Tr}(B_{lj}) = N \text{Tr}(B) + N \sum_{l,j=1, l \neq j}^{2^{k-i}} \text{Tr}(B_{lj}). \end{aligned}$$

Επομένως, χρησιμοποιώντας την συμμετρία του πίνακα B έχουμε

$$\frac{1}{N} \text{Tr}(X^T B X) = \text{Tr}(B) + 2 \sum_{l,j=1, l < j}^{2^{k-i}} \text{Tr}(B_{lj})$$

και επομένως το σφάλμα της εκτίμησης ισούται με $2 \sum_{l,j=1, l < j}^{2^{k-i}} \text{Tr}(B_{lj})$.

(ii) Αν $p \neq 2^k$, καθώς $p \equiv r \pmod{2^i}$, $r \in \mathbb{N}$, $0 \leq r < 2^i$, τότε υπάρχει μία σταθερά $t \in \mathbb{N}$ τέτοια ώστε $p = t2^i + r$. Επομένως, από το Λήμμα 5 έχουμε ότι $X = [H_i, H_i, \dots, H_i, R_i]^T$, όπου $R_i = H_i(1:r, :)$ είναι ο υποπίνακας του

H_i που περιέχει μόνο τις πρώτες r γραμμές του. Στον sample matrix X υπάρχουν Nt πίνακες H_i . Έστω

$$B = \begin{bmatrix} B_{1,1} & \cdots & B_{1,Nt} & \tilde{B}_1^T \\ B_{2,1} & \cdots & B_{2,Nt} & \tilde{B}_2^T \\ \vdots & \vdots & \vdots & \vdots \\ B_{Nt,1} & \cdots & B_{Nt,Nt} & \tilde{B}_{Nt}^T \\ \tilde{B}_1 & \cdots & \tilde{B}_{Nt} & \hat{B}_r \end{bmatrix},$$

όπου τα B_{lj} , $l, j = 1, \dots, Nt$ είναι $N \times N$ blocks, τα \tilde{B}_j , $j = 1, \dots, Nt$ είναι $r \times N$ blocks και το \hat{B}_r είναι ένα $r \times r$ block του πίνακα B . Έχουμε

$$\begin{aligned} \text{Tr}(X^T B X) &= \sum_{l,j=1}^{Nt} \text{Tr}(H_i^T B_{lj} H_i) + \sum_{j=1}^{Nt} \text{Tr}(R_i^T \tilde{B}_j H_i) + \sum_{j=1}^{Nt} \text{Tr}(H_i^T \tilde{B}_j^T R_i) + \text{Tr}(R_i^T \hat{B}_r R_i) = \\ &= \sum_{l,j=1}^{Nt} \text{Tr}(B_{lj} H_i H_i^T) + \sum_{j=1}^{Nt} \text{Tr}(\tilde{B}_j H_i R_i^T) + \sum_{j=1}^{Nt} \text{Tr}(\tilde{B}_j^T R_i H_i^T) + \text{Tr}(\hat{B}_r R_i R_i^T) = \\ &= N \sum_{l,j=1}^{Nt} \text{Tr}(B_{lj}) + 2N \sum_{j=1}^{Nt} \text{Tr}(\tilde{B}_j(1:r, 1:r)) + N \text{Tr}(\hat{B}_r) = \\ &= N \text{Tr}(B) + 2N \sum_{l,j=1, l < j}^{Nt} \text{Tr}(B_{lj}) + 2N \sum_{j=1}^{Nt} \text{Tr}(\tilde{B}_j(1:r, 1:r)). \end{aligned}$$

Επομένως

$$\frac{1}{N} \text{Tr}(X^T B X) = \text{Tr}(B) + 2 \sum_{l,j=1, l < j}^{Nt} \text{Tr}(B_{lj}) + 2 \sum_{j=1}^{Nt} \text{Tr}(\tilde{B}_j(1:r, 1:r))$$

και άρα το σφάλμα της εκτίμησης ισούται με

$$2 \sum_{l,j=1, l < j}^{Nt} \text{Tr}(B_{lj}) + 2 \sum_{j=1}^{Nt} \text{Tr}(\tilde{B}_j(1:r, 1:r)),$$

όπου $\tilde{B}_j(1:r, 1:r)$ αντιπροσωπεύει τον κύριο $r \times r$ υποπίνακα του πίνακα \tilde{B}_j . ■

Παρατήρηση 4. Λαμβάνοντας υπόψη τον συμβολισμό που χρησιμοποιήθηκε στα παραπάνω αποτελέσματα αυτής της ενότητας, από το Λήμμα 5 και το Θεώρημα 3 έπεται ότι μπορούμε να αποφύγουμε την αποθήκευση ενός μεγάλου Sylvester Hadamard πίνακα H_k . Η επανάληψη των Sylvester Hadamard πινάκων H_i , $i < k$ στον *sample matrix* X υποδεικνύει ότι χρειάζεται να αποθηκευτεί μόνο αυτός ο πίνακας. Αυτό επιτυγχάνεται επειδή χρησιμοποιούμε τον κύριο $p \times N$ υποπίνακα του πίνακα H_k , δηλαδή τις πρώτες N διαδοχικές στήλες του.

2.3 Δειγματοληψία από Hadamard πίνακες μέσω doubling construction

Είδαμε ότι οι πίνακες Sylvester Hadamard κατασκευάζονται σύμφωνα με το μοτίβο

$$H_{k+1} = \begin{pmatrix} H_k & H_k \\ H_k & -H_k \end{pmatrix}$$

χρησιμοποιώντας σαν αρχικό πίνακα (ή αλλιώς *seed matrix*) τον 2×2 πίνακα $H_1 = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$, όπως περιγράφεται και στην ενότητα 2.1. Από αυτήν την κατασκευή παίρνουμε πίνακες Hadamard των οποίων η διάσταση είναι πάντα δύναμη του δύο. Μπορούμε επίσης να πάρουμε πίνακες Hadamard διαφορετικών διαστάσεων, η οποία είναι πάντα πολλαπλάσιο του τέσσερα. Αυτό επιτυγχάνεται μέσω doubling construction, δηλαδή χρησιμοποιώντας σαν αρχικό πίνακα στο παραπάνω μοτίβο έναν πίνακα Hadamard διάστασης διαφορετικής του 2. Είναι φανερό ότι οι πίνακες Sylvester Hadamard είναι μια ειδική περίπτωση του doubling construction. Το παρακάτω Θεώρημα γενικεύει το Θεώρημα 3 και μπορεί να αποδειχθεί με παρόμοιο τρόπο.

Θεώρημα 4.

Εστω $A \in \mathbb{R}^{p \times p}$ ένας συμμετρικός πίνακας, h_s η διάσταση ενός *seed* Ha-

damard πίνακα που υποθέτουμε ότι υπάρχει και $k \in \mathbb{N}$. Αν θεωρήσουμε ως *sample matrix* X τον $p \times N$ κύριο υποπίνακα του πίνακα *Hadamard* διάστασης $2^k h_s$, όπου $N = 2^i h_s$, $i < k$, τότε ο πίνακας XX^T είναι αραιός με μονάδες στην κύρια διαγώνιο και σε κάθε $\kappa 2^i$ υπερ και υποδιαγώνιο, $\forall \kappa \in \mathbb{N}$ τέτοιο ώστε $\kappa 2^i \leq p$.

2.4 Εκτιμήσεις χρησιμοποιώντας άλλους σχεδιασμούς

Το Πρόρισμα 5 δηλώνει ότι το σφάλμα της εκτίμησης επηρεάζεται μόνο από τα μη διαγώνια στοιχεία του πίνακα A . Επομένως, μπορούμε εναλλακτικά να χρησιμοποιήσουμε άλλους σχεδιασμούς X οι οποίοι να ελαχιστοποιούν το πλήθος των μη μηδενικών στοιχείων εκτός της κύρια διαγωνίου του πίνακα S . Στη συνέχεια, επεκτείνουμε το αποτέλεσμα της Πρότασης 4 για δείγματα που αποτελούνται από διανύσματα με στοιχεία 0, 1 και -1 .

Πρόταση 5.

Έστω $f(A) \in \mathbb{R}^{p \times p}$ ένας συμμετρικός πίνακας με $\text{Tr}(f(A)) \neq 0$. Έστω $x \in \mathbb{R}^p$ ένα διάνυσμα του οποίου τα στοιχεία παίρνουν τις τιμές 0, με πιθανότητα ρ , και 1, -1 με ίση πιθανότητα (έστω X μία διακριτή τυχαία μεταβλητή που παίρνει τιμές 1 και -1 με ίση πιθανότητα και 0 με πιθανότητα ρ και έστω x ένα διάνυσμα με p ανεξάρτητα δείγματα από την X). Τότε

$$E((x, f(A)x)) = (1 - \rho) \text{Tr}(f(A)).$$

Απόδειξη

Έστω $x = (x_1, \dots, x_n)^T$ ένα διάνυσμα με p ανεξάρτητα δείγματα από την X . Τότε, ισχύει

$$E(x_i x_j) = E(x_i) E(x_j) \quad \text{και} \quad E(x_i) = 0, \quad i, j = 1, \dots, p$$

και επομένως $E(x_i x_j) = E(x_i)E(x_j) = 0$. Επίσης ισχύει ότι $E(x_i^2) = 1 - \rho$. Έχουμε $(x, f(A)x) = \sum_{i,j} (f(A))_{ij} x_i x_j$ που συνεπάγεται

$$E(x, f(A)x) = E\left(\sum_{i,j} (f(A))_{ij} x_i x_j\right) = \sum_{i,j} (f(A))_{ij} E(x_i x_j)$$

δηλαδή

$$E(x, f(A)x) = \sum_i (f(A))_{ii} E(x_i^2) + 2 \sum_{i < j} (f(A))_{ij} E(x_i x_j),$$

καθώς $(f(A))_{ij} = (f(A))_{ji}$ αφού ο πίνακας $f(A)$ είναι συμμετρικός. Επομένως,

$$E(x, f(A)x) = \sum_i (f(A))_{ii} (1-\rho) + 0 = (1-\rho) \sum_i (f(A))_{ii} = (1-\rho) \text{Tr}(f(A)).$$

■

Λήμμα 6.

Έστω $A \in \mathbb{R}^{p \times p}$ ένας συμμετρικός πίνακας και έστω $X = [x_1, \dots, x_N] \in \mathbb{R}^{p \times N}$ ο *sample matrix*, όπου θεωρούμε τα διανύσματα x_i , $i = 1, \dots, N$ όπως περιγράφονται στην Πρόταση 5.

(i) $\text{Tr}(f(A)) \simeq \frac{1}{(1-\rho)N} \sum_{i=1}^N x_i^T f(A) x_i = \text{Tr}(X^T f(A) X) / (N(1-\rho))$.

(ii) Το απόλυτο σφάλμα της προσέγγισης που περιγράφεται στο (i) είναι $\text{Err} = |\text{Tr}(f(A)(I_p - S))|$, όπου $S = X X^T / (N(1-\rho))$.

Απόδειξη

(i) Έπεται άμεσα από την Πρόταση 5 καθώς

$$\left(\sum_{i=1}^N x_i^T f(A) x_i\right) / N \simeq (1-\rho) \text{Tr}(f(A)).$$

(ii) $\text{Err} = \left| \text{Tr}(f(A)) - \frac{\text{Tr}(X^T f(A) X)}{(1-\rho)N} \right| = \left| \text{Tr}(f(A)) - \frac{\text{Tr}(f(A) X X^T)}{(1-\rho)N} \right| =$

$$= |\text{Tr}(f(A)(I_p - S))|, \quad S = \frac{XX^T}{(1 - \rho)N}.$$

■

Μπορούμε να κατασκευάσουμε έναν sample matrix X διάστασης $p \times N$, όπου $N \ll p$, κατά τρόπο παρόμοιο όπως κατασκευάζεται ένα σχεδιασμός CCD, (central composite design) (βλ. [45]). Αυτός ο σχεδιασμός αποτελείται από τρία μέρη:

- Αρχικά, χρησιμοποιείται ένας πίνακας Hadamard, ή μέρος ενός πίνακα Hadamard, που ονομάζεται ως το factorial part του CCD.
- Στην συνέχεια προστίθεται ένα σετ κεντρικών σημείων (central points), που είναι μηδενικά διανύσματα-γραμμές μήκους N .
- Τέλος, θεωρούμε ένα σετ από αξονικά σημεία (axial points), που είναι οι γραμμές του πίνακα E_{2N} διάστασης $2N \times N$, όπου

$$E_{2N} = \begin{pmatrix} \alpha & 0 & 0 & \dots & \dots & 0 \\ -\alpha & 0 & 0 & \dots & \dots & 0 \\ 0 & \alpha & 0 & \dots & \dots & 0 \\ 0 & -\alpha & 0 & \dots & \dots & 0 \\ \dots & & & & & \dots \\ 0 & 0 & 0 & 0 & \dots & \alpha \\ 0 & 0 & 0 & 0 & \dots & -\alpha \end{pmatrix}.$$

Η τιμή της παραμέτρου α καθορίζεται από τον σχεδιαστή. Όμως, κάποιες συγκεκριμένες τιμές αυτής της παραμέτρου μπορούν να δώσουν επιθυμητές ιδιότητες στον σχεδιασμό (βλ. [45]).

Ένας $p \times N$ sample matrix X μπορεί να κατασκευαστεί χρησιμοποιώντας τα προαναφερθέντα πειραματικά σετ, όπου $N = 2^i$, για μία σταθερά $i \in \mathbb{N}$. Ο πίνακας X μπορεί να αποτελείται από έναν ή παραπάνω πίνακες Hadamard διάστασης N , κάποια κεντρικά σημεία και/ή κάποια αξονικά σημεία για $\alpha = 1$.

Χρησιμοποιούμε πάλι μέγεθος δείγματος που είναι δύναμη του 2, ώστε να επωφεληθούμε της ορθογωνιότητας των γραμμών των πινάκων Hadamard στον sample matrix X . Η συμπεριφορά των πινάκων Hadamard μελετήθηκε στην Ενότητα 2.2. Είδαμε ότι τα μη μηδενικά στοιχεία εκτός της κυρίας διαγωνίου στον πίνακα XX^T επηρεάζουν το σφάλμα της εκτίμησης.

Αν χρησιμοποιήσουμε τον $N \times tN$ κύριο υποπίνακα ενός Hadamard πίνακα, όπου $t \in \mathbb{N}$, $tN < p$, και τον επαυξήσουμε με $p - tN$ μηδενικά διανύσματα-γραμμές μήκους N , κατασκευάζοντας κατ' αυτόν τον τρόπο έναν sample matrix X , παρατηρούμε ότι έχουμε λιγότερα μη μηδενικά στοιχεία εκτός της διαγωνίου στον πίνακα XX^T . Αυτό συνεπάγεται επίσης ότι $p - tN$ διαγώνια στοιχεία του πίνακα XX^T γίνονται ίσα με μηδέν.

Αν το p είναι δύναμη του δύο, τότε το πλήθος των κεντρικών σημείων στον sample matrix X θα πρέπει να είναι επίσης μία δύναμη του δύο για μία αποτελεσματική εκτίμηση. Παρατηρούμε ότι εάν $N = p/2$, χρησιμοποιώντας έναν sample matrix X που αποτελείται από έναν πίνακα Hadamard διάστασης N και από N κεντρικά σημεία, τότε δεν υπάρχουν μη μηδενικά στοιχεία εκτός της διαγωνίου στον πίνακα XX^T . Αυτό μπορεί να μειώσει σημαντικά το σφάλμα της εκτίμησης, αναλόγως βέβαια και τον πίνακα A . Για παράδειγμα αυτός ο σχεδιασμός δουλεύει καλά σε πίνακες των οποίων τα διαγώνια στοιχεία είναι της ίδιας τάξης ή οι τιμές τους είναι πολύ κοντά η μία στην άλλη.

Επι προσθέτως, κάποιες γραμμές του πίνακα E_{2N} (αξονικά σημεία) μπορούν να χρησιμοποιηθούν για $\alpha = 1$, σε συνδυασμό με τους πίνακες Hadamard και τα κεντρικά σημεία. Σε αυτήν την περίπτωση, οι στήλες του πίνακα E_{2N} είναι ακόμα γραμμικά ανεξάρτητες. Χρησιμοποιώντας αυτά τα αξονικά σημεία στον sample matrix X , συνεπάγεται ότι κάποια διαγώνια στοιχεία του XX^T παίρνουν τιμή ίση με 1 αντί για N . Επιπλέον, έχουμε περισσότερα μη μηδενικά στοιχεία εκτός της διαγωνίου στον πίνακα XX^T , συγκριτικά με τον sample matrix που περιέχει μόνο τους πίνακες Hadamard και κάποια κεντρικά σημεία. Παρ' όλα αυτά, αυτές οι τιμές είναι ίσες με 1 ή -1 , που σημαίνει

ότι η αντίστοιχη είσοδος του πίνακα A έχει μικρή επίδραση στο σφάλμα της εκτίμησης.

Άλλοι σχεδιασμοί, βασισμένοι στα BBDs, (Box-Behnken designs) (βλ. [14]) μπορούν επίσης να μελετηθούν και να εφαρμοστούν κατά αντίστοιχο τρόπο όπως τα CCDs, (central composite designs).

Κεφάλαιο 3

Αριθμητικά παραδείγματα

Στη συνέχεια, εξετάζουμε τους στοχαστικούς εκτιμητές για την εκτίμηση του ίχνους συναρτήσεων πινάκων, για διάφορες συναρτήσεις πινάκων, χρησιμοποιώντας ως sample matrix X ένα μέρος ενός πίνακα Hadamard, αντί για διανύσματα με εισόδους i.i.d. Rademacher τυχαίες μεταβλητές. Θα αναφερόμαστε σε αυτήν την μέθοδο ως Hadamard sampling. Στα παρακάτω παραδείγματα, αυτή η μέθοδος συγκρίνεται με την δειγματοληψία Rademacher (Rademacher sampling).

Επίσης, εξετάζουμε την συμπεριφορά των στοχαστικών εκτιμητών που υπολογίζονται χρησιμοποιώντας άλλους sample matrices X οι οποίοι έχουν τιμές 0, 1 και -1 .

Για τα αποτελέσματα που παρουσιάζονται σε αυτό το Κεφάλαιο, χρησιμοποιήσαμε την μέθοδο της παρεκβολής και τις εκτιμήσεις που παρουσιάζονται στις εργασίες [28], [29], [17], [16] για την εκτίμηση των τετραγωνικών μορφών $x_i^T f(A)x_i$.

Για την κατασκευή των απαιτούμενων διανυσμάτων x_i στην μέθοδο Rademacher sampling, χρησιμοποιήσαμε τον uniform generator τυχαίων αριθμών μεταξύ 0 και 1 του MATLAB. Αν ο τυχαίος αριθμός που παίρναμε ήταν μικρότερος ή ίσος του 0.5, η αντίστοιχη είσοδος του x_i έπαιρνε την τιμή -1 ενώ

αν ήταν μεγαλύτερος του 0.5 και μικρότερος ή ίσος του 1, τότε η αντίστοιχη είσοδος του x_i γινόταν ίση με +1.

Το σχετικό σφάλμα ορίζεται ως το πηλίκο της απόλυτης τιμής της διαφοράς της ακριβούς τιμής του ίχνους από τον εκτιμητή της, προς την απόλυτη τιμή της ακριβούς τιμής του ίχνους. Γενικά είναι το πηλίκο της απόλυτης τιμής της διαφοράς της θεωρητικής τιμής μιας ποσότητας από τον εκτιμητή της, προς την απόλυτη τιμή της θεωρητικής αυτής τιμής. Στην εργασία [56], προτείνεται η διασπορά ως μέτρο σύγκλισης στην ακριβή τιμή του ίχνους.

Όλοι οι υπολογισμοί εκτελέστηκαν στο MATLAB (R2011B), 64-BIT, σε έναν Intel Core i7 υπολογιστή, με 8 Gb RAM.

Παράδειγμα 1

Εξετάσαμε τον καλής κατάστασης πίνακα Kms με παράμετρο $w = 0.2$ και τον αραιό, κακής κατάστασης πίνακα $Poisson$, που δίνονται χρησιμοποιώντας την εσωτερική συνάρτηση *gallery* του MATLAB. Επίσης, εξετάσαμε έναν ψευδοαραιό πίνακα Q με εισόδους $q_{ij} = e^{-2|i-j|}$.

Στους Πίνακες 3.1, 3.2, 3.3 και 3.4 συγκρίνουμε την συμπεριφορά των μεθόδων Hadamard και Rademacher sampling στην εκτίμηση των ποσοτήτων $\text{Tr}(A^r)$, $\text{Tr}(\sqrt{A})$, $\text{Tr}(\log A)$ και $\text{Tr}(\exp A)$, χρησιμοποιώντας δείγμα μεγέθους $N = 32$. Αξίζει να σημειωθεί ότι αν χρησιμοποιήσουμε πλήθος στηλών ενός πίνακα Hadamard, σαν δείγμα, που δεν είναι δύναμη του δύο, τότε δεν επωφελούμαστε των ιδιοτήτων των πινάκων Hadamard.

Η διάσταση των πινάκων Kms και Q είναι 1024, ενώ του πίνακα $Poisson$ είναι 625, εκτός και αν μία διαφορετική διάσταση αναγράφεται εντός της παρενθέσεως δίπλα στο όνομα του πίνακα.

Η μέθοδος Hadamard sampling εξαρτάται από την δομή του πίνακα. Τα μεγαλύτερα σε μέγεθος στοιχεία των παραπάνω πινάκων που εξετάστηκαν συσσωρεύονται κοντά στην διαγώνιο. Σε πίνακες με αυτήν την ειδική δομή, οι μη επιθυμητές υπερ και υποδιαγώνιοι του πίνακα XX^T επηρεάζουν λιγότερο

Matrix	r	Hadamard	Radem
Kms	5	2.2262e-3	1.8437e-2
Kms	3	1.6599e-13	5.6381e-4
Kms	-1	6.8386e-4	1.7821e-3
Kms (1000)	-1	6.8390e-4	4.0573e-3
Q	5	5.8103e-4	4.6548e-3
Q	3	2.2226e-13	1.1568e-3
Q	-1	1.4342e-4	2.0664e-3
Q (1000)	-1	1.4343e-4	2.3528e-3

Πίνακας 3.1: Σχετικά σφάλματα εκτιμώντας το $\text{Tr}(A^r)$.

Matrix	Hadamard	Radem
Kms	2.9004e-5	2.7576e-4
Kms (1000)	2.9006e-5	6.7120e-4
Poisson	1.8306e-4	5.4810e-3
Q	5.8258e-6	1.7984e-4
Q (1000)	5.8262e-6	4.0502e-4

Πίνακας 3.2: Σχετικά σφάλματα εκτιμώντας το $\text{Tr}(\sqrt{A})$.

Matrix	Hadamard	Radem
KMS	4.4986e-3	4.0812e-2
KMS (1000)	4.4990e-3	1.8449e-2
Poisson	5.9182e-3	5.8153e-2
Q	2.0072e-3	5.0351e-2
Q (1000)	2.0074e-3	1.3521e-2

Πίνακας 3.3: Σχετικά σφάλματα εκτιμώντας το $\text{Tr}(\log A)$.

Matrix	Hadamard	Radem
KMS	3.5899e-5	1.3588e-3
KMS (1000)	9.7030e-5	1.4421e-3
Poisson	1.6063e-4	1.3218e-2
Q	6.7043e-6	2.3351e-4
Q (1000)	1.9173e-5	2.7839e-3

Πίνακας 3.4: Σχετικά σφάλματα εκτιμώντας το $\text{Tr}(expA)$.

το σφάλμα της εκτίμησης και επομένως η Hadamard sampling βελτιώνει τα αποτελέσματα. Όπως είδαμε, μπορούμε να επιτύχουμε καλύτερες εκτιμήσεις αν το πλήθος των στηλών που θα χρησιμοποιήσουμε στο δείγμα είναι δύναμη του δύο. Για παράδειγμα, χρησιμοποιήσαμε 33 στήλες και παρατηρήσαμε ότι το σχετικό σφάλμα είναι χειρότερο απ' ό,τι αν χρησιμοποιήσουμε 32 στήλες στην Hadamard sampling, ενώ είναι συγκρίσιμο ή χειρότερο με εκείνο που παίρνουμε χρησιμοποιώντας την Rademacher sampling.

Στη συνέχεια, στα Γραφήματα 3.1 και 3.2, βλέπουμε το σχετικό σφάλμα που προκύπτει εκτιμώντας το ίχνος $\text{Tr}(A^3)$ του πίνακα *kms* (πάνω) και του πίνακα *Q* (κάτω), όσο το πλήθος N των διανυσμάτων του δείγματος αυξάνεται, χρησιμοποιώντας είτε Rademacher τυχαία διανύσματα (Γράφημα 3.1) είτε στήλες ενός πίνακα Hadamard (Γράφημα 3.2).

Στο Γράφημα 3.2 παρατηρούμε ότι το σχετικό σφάλμα σταδιακά μειώνεται όσο το μέγεθος του δείγματος αυξάνεται στην Hadamard sampling. Από την άλλη μεριά, στο Γράφημα 3.1 παρατηρούμε ότι η Rademacher sampling δεν έχει την ίδια σταθερή συμπεριφορά.

Παράδειγμα 2

Στη συνέχεια, παρουσιάζονται δύο πραγματικές εφαρμογές που χρειάζονται τον υπολογισμό του ίχνους $\text{Tr}(f(A))$, και προέρχονται από την ανάλυση δικτύων και την φυσική στερεάς κατάστασης (solid state physics).

Στην ανάλυση δικτύων, ο δείκτης Estrada (βλ. [24]) ορίζεται ως το ίχνος $\text{Tr}(\exp(A))$, για έναν πίνακα γειτνίασης A . Θεωρούμε τον πίνακα γειτνίασης *erdrey* διάστασης p , ο οποίος αντιπροσωπεύει απλά συνδεδεμένα γραφήματα και προέρχεται από το toolbox CONTEST του MATLAB (βλ. [60]).

Στον Πίνακα 3.5 συγκρίνουμε το σχετικό σφάλμα για την εκτίμηση τον δείκτη Estrada του πίνακα *erdrey* για διάφορες διαστάσεις p , χρησιμοποιώντας Hadamard και Rademacher sampling.

p	Hadamard	Radem
500	3.5848e-2	1.3520e-1
1000	1.5316e-2	1.2827e-1
2000	6.8972e-3	1.4541e-1

Πίνακας 3.5: Σχετικά σφάλματα εκτιμώντας τον δείκτη Estrada.

Στη φυσική στερεάς κατάστασης (solid state physics) ο υπολογισμός της συνολικής ενέργειας της ηλεκτρονικής δομής απαιτεί τον υπολογισμό των αθροισμάτων μερικών ιδιοτιμών $\sum_{k=1}^m \lambda_k$, $m < p$. Έστω $A \in \mathbb{R}^{p \times p}$ ένας συμμετρικός πίνακας διάστασης p και $\lambda_1, \lambda_2, \dots, \lambda_p \in \mathbb{R}$ οι ιδιοτιμές του. Έστω μ μία πραγματική μεταβλητή τέτοια ώστε $\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_m \leq \mu \leq \lambda_{m+1} \leq \dots \leq \lambda_p$. Ισχύει $\sum_{k=1}^m \lambda_k \simeq \text{Tr}(f(A))$ για $f(z) = z/(1 + \exp((z - \mu)/c))$, c μία σταθερά (βλ. [4]).

Στον Πίνακα 3.6, εξετάσαμε τον πίνακα *Poisson* διάστασης p , χρησιμοποιώντας Hadamard και Rademacher sampling.

Στους Πίνακες 3.5 και 3.6 θεωρήσαμε δύο περιπτώσεις στις οποίες η εκτίμηση που δοκιμάστηκε για την ποσότητα $x_i^T f(A) x_i$ δεν έχει πολύ καλή συμπεριφορά χρησιμοποιώντας την Rademacher sampling. Παρατηρούμε όμως ότι η Hadamard sampling μπορεί να δώσει καλύτερη προσέγγιση σε αυτές τις περιπτώσεις.

p	Hadamard	Radem
625	4.5106e-2	1.1132e-1
1225	7.1315e-3	1.2817e-1
2704	3.9283e-2	1.3245e-1

Πίνακας 3.6: Σχετικά σφάλματα εκτιμώντας το μερικό άθροισμα των ιδιοτιμών ($c = 1$, $\mu = 5$).

Παράδειγμα 3

Θεωρούμε τον nearly orthogonal Chebyshev–Vandermonde–like *orthog* πίνακα A του οποίου τα στοιχεία είναι $a_{ij} = \cos((i-1)(j-1)\pi/(n-1))$. Θεωρούμε τον πίνακα $G = AA^T$. Οι πίνακες G είναι καλής κατάστασης και συμμετρικοί, θετικά ορισμένοι. Επίσης θεωρούμε τον μη συμμετρικό πίνακα *Parter* B του οποίου τα στοιχεία είναι $b_{ij} = 1/(i-j+0.5)$. Ο B είναι Cauchy και Toeplitz πίνακας. Θέτουμε $T = B^T B$ ώστε ο T να είναι συμμετρικός. Θεωρούμε τον *Prolate* πίνακα P ο οποίος είναι συμμετρικός και Toeplitz με $y = 0.9$. Θεωρούμε επίσης τον πίνακα *Kms* K με παράμετρο $w = 0.2$. Οι παραπάνω πίνακες προέρχονται καλώντας την εσωτερική συνάρτηση *gallery* του MATLAB. Επιπλέον, εξετάζουμε ένα παράδειγμα που προτείνεται από τον Strakoš στην εργασία [57], το οποίο εξετάζεται επίσης στο [32]. Έστω Λ ένας διαγώνιος πίνακας με στοιχεία $\lambda_{ii} = \lambda_1 + \left(\frac{i-1}{p-1}\right)(\lambda_p - \lambda_1)d^{p-i}$, $i = 1, \dots, p$, και έστω V ένας ορθογώνιος πίνακας των ιδιοδιανυσμάτων του τριδιαγώνιου πίνακα $(-1, 2, -1)$. Θεωρούμε τον πίνακα $S = V^T \Lambda V$. Εξετάσαμε αυτόν τον πίνακα για $\lambda_1 = 0.1$, $\lambda_p = 100$ και $d = 0.9$. Τέλος, θεωρούμε έναν τυχαίο πίνακα R_1 , που δίνεται χρησιμοποιώντας τη συνάρτηση *rand* του MATLAB και θέτουμε $R = R_1 R_1^T$.

Εξετάσαμε τους παραπάνω πίνακες διάστασης 64, χρησιμοποιώντας δείγμα μεγέθους $N = 32$. Στον Πίνακα 3.7 συγκρίνουμε το σχετικό σφάλμα εκτιμώντας το $\text{Tr}(A^3)$ χρησιμοποιώντας Rademacher τυχαίες μεταβλητές, τις πρώτες 32

στήλες του Sylvester Hadamard πίνακα διάστασης 64 και έναν σχεδιασμό $X = [H_5; Z_{32}]$, όπου H_5 είναι ο Sylvester Hadamard πίνακας διάστασης 32 και Z_{32} είναι ένας 32×32 πίνακας με όλες του τις εισόδους ίσες με μηδέν.

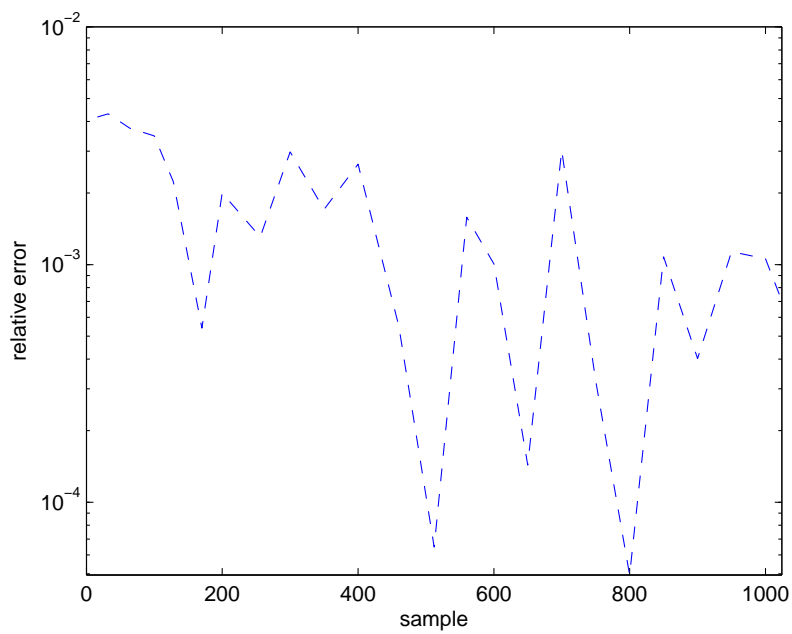
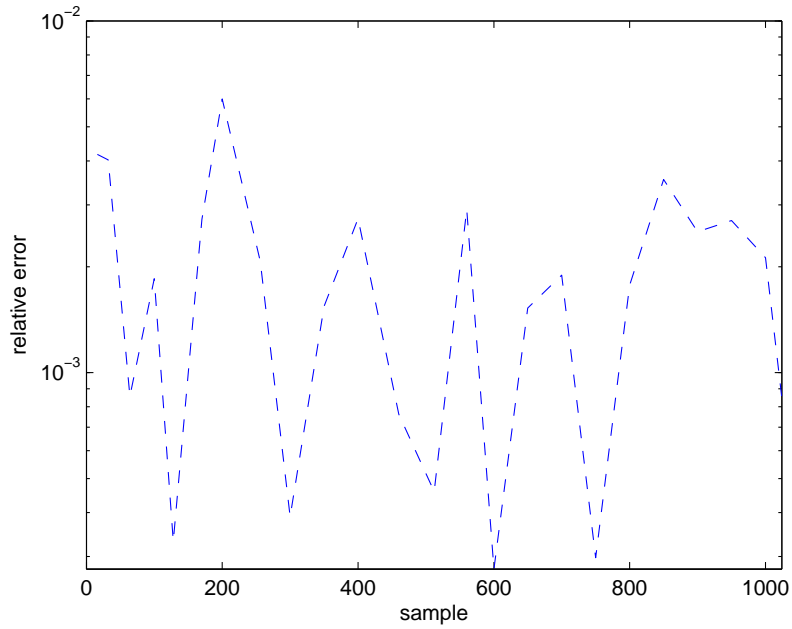
Matrix	Hadamard	Radem	X
G	1.6195e-1	1.0635e-1	1.5994e-16
P	1.0921e-2	2.6957e-2	0
K	1.5734e-14	1.2056e-2	1.7679e-16
T	1.0935e-2	1.3388e-2	7.9862e-3
S	2.1685e-3	1.5096e-1	2.8508e-16
R	9.9573e-1	9.9601e-1	5.8170e-3

Πίνακας 3.7: Σχετικά σφάλματα εκτιμώντας το $\text{Tr}(A^3)$.

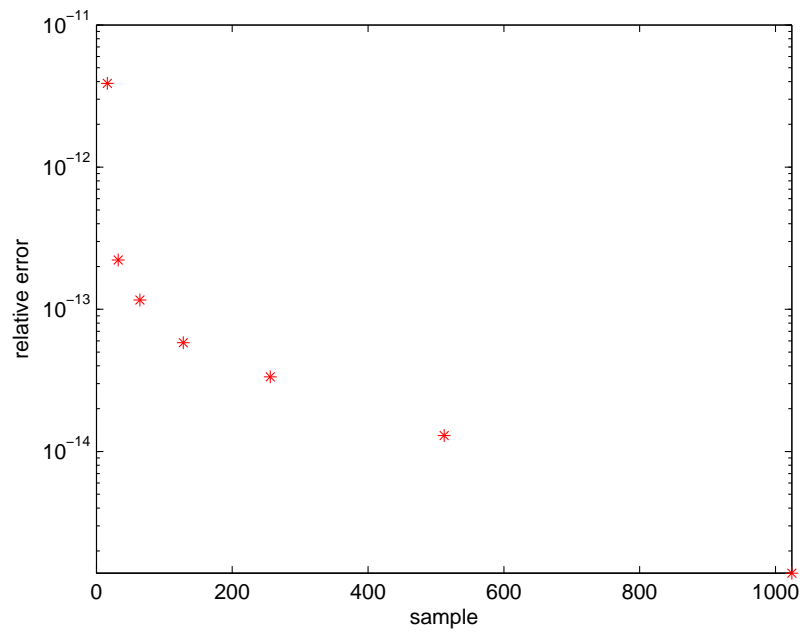
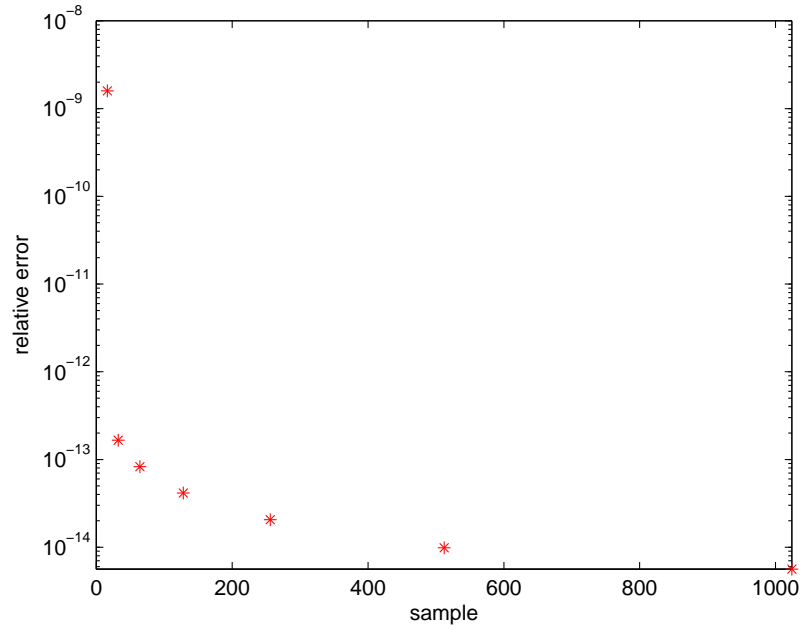
Εξετάζουμε έναν τυχαίο πίνακα R_2 διάστασης 128, που δίνεται από τη συνάρτηση *rand* του MATLAB, και θέτουμε $R_n = R_2^T R_2$. Στον Πίνακα 3.8 συγκρίνουμε το σχετικό σφάλμα εκτιμώντας το $\text{Tr}(A^3)$ χρησιμοποιώντας 32 διανύσματα Rademacher, τις πρώτες 32 στήλες του Sylvester Hadamard πίνακα διάστασης 128 και έναν σχεδιασμό $X = [H_5; E_{64}; Z_{32}]$, όπου H_5 είναι ο Sylvester Hadamard πίνακας διάστασης 32, E_{64} είναι ο πίνακας με τα αξονικά σημεία και Z_{32} είναι ένας 32×32 πίνακας με όλα του τα στοιχεία μηδενικά. Παρατηρούμε ότι αυτός ο πίνακας έχει κακή συμπεριφορά, ακόμα και αν χρησιμοποιήσουμε Hadamard sampling. Παρ'όλα αυτά, η χρήση των αξονικών σημείων και κάποιων κεντρικών σημείων στον sample matrix X μπορεί να οδηγήσει σε μια καλή εκτίμηση του $\text{Tr}(A^3)$.

Matrix	Hadamard	Radem	X
R_n	2.9932e0	9.9672e-1	6.0443e-2

Πίνακας 3.8: Σχετικά σφάλματα εκτιμώντας το $\text{Tr}(A^3)$.



Σχήμα 3.1: Σχετικό σφάλμα εκτιμώντας το $\text{Tr}(A^3)$, όσο αυξάνεται το πλήθος των τυχαίων Rademacher διανυσμάτων στο δείγμα.



Σχήμα 3.2: Σχετικό σφάλμα εκτιμώντας το $\text{Tr}(A^3)$, όσο αυξάνεται το πλήθος των στηλών του πίνακα Hadamard στο δείγμα.

Βιβλιογραφία

- [1] H. AVRON, *Counting triangles in large graphs using randomized matrix trace estimation*, Proceedings of KDD–LDMTA'10, July 2010.
- [2] H. AVRON, S. TOLEDO, *Randomized algorithms for estimating the trace of an implicit symmetric positive semi-definite matrix*, Journal of the ACM, 58(2), 2011.
- [3] Z.-Z. BAI, M. FAHEY, G. GOLUB, *Some large scale computation problems*, J. Comput. Appl. Math., 74 (1996), pp. 71–89.
- [4] Z. BAI, M. FAHEY, G. GOLUB, M. MENON, E. Richter, *Computing partial eigenvalue sum in electronic structure calculation*, Report SCCM-98–03, Stanford University (1998), 19 pp.
- [5] Z. BAI, G.H. GOLUB, *Bounds for the trace of the inverse and the determinant of symmetric positive definite matrices*, Annals Numer. Math., 4 (1997), pp. 29–38.
- [6] C. BERNARDI, Y. MADAY, Spectral methods, in *Handbook of Numerical Analysis*, Vol. V, P.G. Ciarlet and J.L. Lions eds., North–Holland, Amsterdam, 1997, pp. 209–485.
- [7] M. BERTERO, P. BOCCACCI, *Introduction to Inverse Problems in Imaging*, Institute of Physics Publishing, Bristol, Philadelphia, 1998.

- [8] C. BEKAS, A. CURIONI, I. FEDULOVA, *Low-cost data uncertainty quantification*, *Concurr. Comput.: Practise and Experience*, 24 (2012), pp. 908–920.
- [9] C. BEKAS, E. KOKIOPOULOU, Y. SAAD, *An estimator for the diagonal of a matrix*, *Applied Numerical Mathematics*, 57 (2007), pp. 1214–1229.
- [10] M. BENZI, C. KLYMKO, *Total Communicability as a centrality measure*, *Journal of Complex Networks*, 1(2), (2013), pp. 124–149.
- [11] M. BENZI, P. BOITO, *Quadrature rule-based bounds for functions of adjacency matrices*, *Linear Algebra and its Applications*, 433 (2010), pp. 637–652.
- [12] M. BELLALIJ, L. REICHEL, G. RODRIGUEZ, H. SADOK, *Bounding matrix functionals via partial global block Lanczos decomposition*, *Applied Numerical Mathematics*, 94, (2015), pp. 127-139.
- [13] M. BENZI, N. RAZOUK, *Decay bounds and $O(N)$ algorithms for approximating functions of sparse matrices*, *Elec. Trans. Numer. Anal.*, 28 (2007), pp. 16–39.
- [14] G. BOX, D. BEHNKEN, *Some new three level designs for the study of quantitative variables*, *Technometrics*, 2 (1960), pp. 455–475.
- [15] C. BREZINSKI, *Error estimates for the solution of linear systems*, *SIAM J. Sci. Comput.*, 21 (1999), pp. 764–781.
- [16] C. BREZINSKI, P. FIKA, M. MITROULI, *Estimations of the trace of powers of positive self-adjoint operators by extrapolation of the moments*, *Elec. Trans. Numer. Anal.*, 39 (2012), pp. 144–155.
- [17] C. BREZINSKI, P. FIKA, M. MITROULI, *Moments of a linear operator on a Hilbert space, with applications to the trace of the inverse of*

-
- matrices and the solution of equations*, Numer. Linear Algebra Appl., 19 (2012), pp. 937–953.
- [18] C. BREZINSKI, M. RAYDAN, *Cauchy–Schwartz and Kantorovich type inequalities for scalar and matrix moment sequences*, Adv. Comput. Math., 26 (2007), pp. 71–80.
- [19] C. BREZINSKI, G. RODRIGUEZ, S. SEATZU, *Error estimates for linear systems with applications to regularization*, Numer. Algorithms, 49 (2008), pp. 85–104.
- [20] C. BREZINSKI, G. RODRIGUEZ, S. SEATZU, *Error estimates for the regularization of least squares problems*, Numer. Algorithms, 51 (2009), pp. 61–76.
- [21] B. CERANKA, M. GRACZYL, *A–optimal chemical balance weighing design with correlated errors*, J. Appl. Math. & Computing, 16 (2004), pp. 143–150.
- [22] B. N. DATTA, *Numerical Linear Algebra and Applications*, Brooks/Cole Publishing Company, Pacific Grove, 1995.
- [23] S–J. DONG, K–F. LIU, *Stochastic estimation with Z_2 noise*, Phys. Lett., B328 (1994), pp. 130–136.
- [24] E. ESTRADA, D. J. HIGHAM, *Network properties revealed through matrix functions*, University of Strathclyde Mathematics Research Report 17, 2008.
- [25] E. ESTRADA, D. J. HIGHAM, *Network properties revealed through matrix functions*, SIAM Review, 52(No. 4), (2010), pp. 696–714.
- [26] C. FENU, D. MARTIN, L. REICHEL, G. RODRIGUEZ, *Network analysis via Partial Spectral Factorization and Gauss Quadrature*, SIAM J. Sci. Comput., 35(4), (2013), pp. 2046–2068.

- [27] P. ΦΙΚΑ, C. ΚΟΥΚΟΥΒΙΝΟΣ, *Stochastic estimators of the trace for functions of matrices via Hadamard matrices*, Communications in Statistics - Simulation and Computation, to appear.
- [28] P. ΦΙΚΑ, M. ΜΙΤΡΟΥΛΙ, *Estimation of the bilinear form $y^*f(A)x$ for Hermitian matrices*, Linear Algebra Appl., to appear.
- [29] P. ΦΙΚΑ, M. ΜΙΤΡΟΥΛΙ, P. ΡΟΥΠΑ, *Estimates for the bilinear form $x^T A^{-1}y$ with applications to linear algebra problems*, Elec. Trans. Numer. Anal., 43 (2014), pp. 70–89.
- [30] D. ΓΙΡΑΡΔ, *Un algorithme simple et rapide pour la validation croisee generalisee sur des problemes de grande taille*, RR 669-M, Grenoble, France: Informatique et Mathématiques Appliquées de Grenoble, 1987.
- [31] C. I. ΓΕΩΡΓΗΙΟΥ, *Spectral Methods for Differential Problems*, Casa Cartii de Stiinta, Cluj–Napoca, 2007.
- [32] G.H. ΓΟΛΟΥΒ, G. ΜΕΥΡΑΝΤ, *Matrices, Moments and Quadrature with Applications*, Princeton University Press, Princeton, 2010.
- [33] G. H. ΓΟΛΟΥΒ, *Matrix computation and the theory of moments*, Proceedings of the International Congress of Mathematicians (1995), pp. 1440–1448.
- [34] G. H. ΓΟΛΟΥΒ, Z. ΣΤΡΑΚΟŠ, *Estimates in quadratic formulas*, Numer. Algorithms, 8 (1994), pp. 241–268.
- [35] N. J. ΗΙΓΧΑΜ, *Functions of Matrices: Theory and Computation*, SIAM, Philadelphia, 2008.
- [36] P. ΗΙΤΤΣΕΝΚΟ, S. ΚΩΑΠΙΕŃ, *On the Rademacher series*, Progress in probability, 35 (1994), pp. 31–36.

-
- [37] K. J. HORADAM, *Hadamard Matrices and Their Applications*, Princeton University Press, Princeton, 2007.
- [38] R. A. HORN, C.R. JOHNSON, *Matrix Analysis*, Cambridge University Press, Cambridge 1985.
- [39] A. S. HOUSEHOLDER, *The Theory of Matrices in Numerical Analysis.*, New YorkQ Blaisdell, 1964.
- [40] M. HUTCHINSON, *A stochastic estimator of the trace of the influence matrix for Laplacian smoothing splines*, Communications in Statistics - Simulation and Computation, 18 (1989), pp. 1059–1076.
- [41] I. IPSEN, D. LEE, *Determinan Approximation*, Technical Report CRSC-TR03-30, Centre for Research in Scientific Computing, 2003.
- [42] V. KALANTZIS, C. BEKAS, A. CURIONI, E. GALLOPOULOS, *Accelerating data uncertainty quantification by solving linear systems with multiple right-hand sides*, Numer. Algor., 62 (2013), pp. 637–653.
- [43] C. KOUKOUVINOS, K. MYLONA, D. SIMOS, *A hybrid SAGA algorithm for the construction of $E(s^2)$ -optimal cyclic supersaturated designs*, J. Statist. Plann. Inference, 139 (2009), pp. 478–485.
- [44] G. MEURANT, P. TICHÝ, *On computing quadrature-based bounds for the A -norm of the error in conjugate gradients*, Numer. Algorithms, 62 (2013), pp. 163–191.
- [45] D. C. MONTGOMERY, *Design and Analysis of Experiments*, 8th Ed., J. Wiley and Sons, New York, 2012.
- [46] G. MEURANT, *Estimates of the trace of the inverse of a symmetric matrix using the modified Chebyshev algorithms*, Numer. Algorithms, 51 (2009), pp. 309–318.

- [47] F. NATTERER, F. WÜBBELING, *Mathematical Methods in Image Reconstruction*, SIAM, Philadelphia, 2001.
- [48] B. ORTNER, A. R. KRAUTER, *Lower bounds for the determinant and the trace of a class of Hermitian matrices*, *Linear Algebra Appl.*, 236 (1996), pp. 147–180.
- [49] V. PAN, *Estimating the extremal eigenvalues of a symmetric matrix*, *Computers Math. Applic.*, 20 (1990), pp. 17–22.
- [50] T. PETRILA, D. TRIF, *Basics of Fluids Mechanics and Introduction to Computational Fluid Dynamics*, Springer Science+Business Media, New York, 2005.
- [51] F. PUKELSHEIM, *Optimal Design of Experiments*, J. Wiley, New York, 1993.
- [52] P. ROBINSON, A. J. WATHEN, *Variational bounds on the entries of the inverse of a matrix*, *IMA Journal of Numerical Analysis*, 12 (1992), pp. 463–486.
- [53] H. RÖDER, R. N. SILVER, D. A. DRABOLD, J. J. DONG, *Kernel polynomial method for a nonorthogonal electronic-structure calculation of amorphous diamond*, *Phys. Rev. B*, 55 (1997), pp 15382–15385.
- [54] P.E. SAYLOR, D .C. SMOLARSKI, *Why Gaussian quadrature in the complex plane?*, *Numer. Algorithms*, 26 (2001), pp. 251–280.
- [55] B. SIMON, *Trace Ideals and their Applications*, 2nd ed., Amer. Math. Soc., Providence, 2005.
- [56] A. STATHOPOULOS, J. LAEUCHLI, K. ORGINOS, *Hierarchical probing for estimating the trace of the matrix inverse on toroidal lattices*, *SIAM J. Sci. Comput.* 35 (2013), pp. S299-S322.

-
- [57] Z. STRAKOŠ, *On the real convergence rate of the conjugate gradient method*, Linear Alg. Appl., 154–156 (1991), pp. 535–549.
- [58] Z. STRAKOŠ, P. TICHÝ, *On efficient numerical approximation of the bilinear form $c^*A^{-1}b$* , SIAM J. Sci. Comput. (SISC), 33 (2011), pp. 565–587.
- [59] J. TANG, Y. SAAD, *A probing method for computing the diagonal of a matrix inverse*, Numer. Linear Algebra Appl., 19 (2012), pp. 485–501.
- [60] A. TAYLOR, D. J. HIGHAM, CONTEST: Toolbox files and documentation. Available at http://www.mathstat.strath.ac.uk/research/groups/numerical_analysis/contest/toolbox.
- [61] R. TÜRKMEN, Z. ULUKÖK, *On the Frobenius condition number of positive definite matrices*, J. Inequal. Appl., Volume 2010 (2010), Article ID 897279, DOI:10.1155/2010/897279.
- [62] C. S. WONG, J. C. MASARO, *A-optimal design matrices $X = (x_{ij})_{N \times n}$ with $x_{ij} = -1, 0, 1$* , Linear and Multilinear Algebra, 15 (1984), pp. 23–46.
- [63] M. N. WONG, F. J. HICKERNELL, K. I. LIU, *Computing the trace of a function of a sparse matrix via Hadamard-like sampling*, Tech Rep. 377(7/04), (2004), Hong Kong Baptist University.
- [64] L. WU, A. STATHOPOULOS, J. LAEUCHLI, V. KALANTZIS, E. GALLOPOULOS, *Estimating the trace of the matrix inverse by interpolating from the diagonal of an approximate inverse*, Technical Report, Report number: WM-CS-2015-02, College of William and Mary, 2015.

- [65] A. V. ZARELUA, *On congruences for the traces of powers of some matrices*, Proc. Steklov Inst. Math., 263 (2008), pp. 78–98.
- [66] F. ZHANG, *Matrix Theory, Basic Results and Techniques*, New York, Springer, 1999.