Rates of Convergence in the Central Limit Theorem for Markov Chains

Your Name Goes Here, Ph.D.
University of Connecticut, 2026

ABSTRACT

Put your abstract here. Try to use as little symbols as possible. This abstract will be typeset for web publishing and various databases (not by you), and symbols don't transfer well.

Rates of Convergence in the Central Limit Theorem for Markov Chains

Your Name Goes Here

M.Sc. Mathematics, Universiteit Antwerpen, Antwerp, Belgium, 1992
M.Sc. Theoretical Physics, Université Libre de Bruxelles, Brussels, Belgium, 1993

A Dissertation
Submitted in Partial Fulfillment of the
Requirements for the Degree of
Doctor of Philosophy
at the
University of Connecticut

2026

APPROVAL PAGE

Doctor of Philosophy Dissertation

Rates of Convergence in the Central Limit Theorem for Markov Chains

Presented by
Your Name Goes Here, M.Sc. Math., M.Sc. Th.Phys.

Major Advisor ______________________________
Richard Bass

Associate Advisor ______________________________
Evarist Giné-Masdéu

Associate Advisor ______________________________
Yung-Sze Choi

University of Connecticut
2026

ACKNOWLEDGEMENTS

At this point, I would like to thank the many people who - directly or indirectly - have contributed to this dissertation. ....

TABLE OF CONTENTS

Introduction

0.1 TO DO

- To write:
  - Document. Fix bibliography
  - Intro. More formulas in intro.
  - Moments. Prove (again, better): Determinacy result for gaussian weighted L2.
  - Moments. Find good reference for terminology and history.
  - GF/M. Define Hermite polynomials, uni- and multi-variate.
  - GF/M. Prove determinacy from Rodrigues formula
  - RT/GRT. Define RT/GRT on distributions (See Helgason or Gelfand). Discuss: do these extended definitions satisfy the same purpose as the GRT? If so then what is the advantage of the GRT?
  - RT/GRT. Slice theorem: connect sufficient conditions to Fubini conditions.
  - RT/GRT. Work on notation for Gaussian measure. In particular w and ω are too similar.
  - RT/GRT. Search for references on G projection moments etc
  - RT/GRT. Expand background, historical context, etc...
  - Moments. Prove: Bounded regions are uniquely determined by moments? Maybe fits in RT/GRT since it can be formulated in terms of RT.
  - Implementation. Output some figures from Mathematica. For example sample points, moment tables, and reconstructions.
  - Example. Try to prove the conjecture (RT Hermite)
  - Split Ch2 in half: Theory and application.
- To research:

- Everywhere (ongoing). Provide examples for $n = 2, 3$ of as many results as we can.
- GRT. Domains of defintion, i.e. $GRT : L^2 \to L^2$ and such.
- GRT. Integration by parts formula for GRT of $\partial f/\partial x_i$ (found dR/dp and thus dGR/dp)
- GRT. Verify the conjecture for $n = 3$
- GRT. Prove the conjecture on GRT of Hermite polynomials.
- Try to derive GRT of monomials from Hermite polynomials.
- Mathematica. Implement Gaussian shape reconstruction method, try some example shapes. Clearly lay out the choices made: Series of expanding images, or compactified $\mathbb{R}^n$? How to complexify points?

0.2 Shape Reconstruction

The shape reconstruction method: In a 2005 paper, Annie Cuyt et. al. [?Cuyt05] proposed a method for shape reconstruction from moments, via Pade approximants to a multidimensional integral transform. Given a set of multivariate moments of some region A in $\mathbb{R}^n$, the method produces a pixel image approximating A.

Suppose $A \subset \mathbb{R}^n$, and let $f(x)$ be its indicator function. We may assume f is measurable and has bounded support, and thus finite moments. The moment sequence gives Taylor coefficients in a neighborhood of zero for a certain holomorphic function with integral representation

$$g(y) = \int_{\mathbb{R}^n} \frac{f(x)}{1 + \langle x, y \rangle} \, dx \approx \sum_{\alpha \in \mathbb{N}_0^n} \binom{|\alpha|}{\alpha} c_\alpha (-y)^\alpha.$$

Thus the function $g(y)$ may be approximated to a certain degree of accuracy depending on the number of available moments.

At the same time g can be approximated by a multivariate quadrature formula

$$\int_{\mathbb{R}^n} \frac{f(x)}{1 + \langle x, y \rangle} \, dx = \sum_i \frac{1}{1 + \langle x_i, y \rangle} f(x_i) = \sum_i w(x_i, y) f(x_i)$$

Where the nodes x_i lie, for example, on a cubic lattice. We can then sample a Pade approximant to g at some sufficinetly large group of points $y = y_j$ forming a linear system of equations, from which we solve for $f(x_i)$. If all goes well we have approximations of $f(x_i)$ on a node lattice, which can be turned into a pixel image of A.

For now we will gloss over the numerical discussion of quadrature and linear systems, taking for granted that such methods are applicable in at least some simple cases. Computational tests (to be included in, say, section 10) further support the validity of the method. Our focus will be on demonstrating that one can approximate g by rational Padé approximants constructed from moments. To this end we show that, when restricted to one dimensional subspaces, g is equivalent to the Stieltjes transform of the Radon transform of f at a fixed projection angle.

$$g(z\omega) = \int_{-\infty}^{\infty} \frac{R_f(\omega, p)}{1 + zp} dp$$

where $\omega \in S^{n-1}, z \in \mathbb{R}$. Thus we are able to leverage the more well understood properties of these univariate transforms.

In particular, the "projection moments" of $R_f(\omega, p)$ in any fixed direction ω, may be computed from the multivariate moments of f. Further, it can be shown that Padé approximants to a moment sequence converge to the Stieltjes transform, and thus to g. Returning to shape reconstruction, it then follows that one can approximate g at various sample points $y_j = z_j\omega_j$ via Padé approximants on linear subspaces at a selection of angles ω, as is needed to form our linear system.

Now let us briefly discuss our proposed modification to the shape reconstruction method, the theoretical complications and computational considerations. Essentially what we propose is to recreate the method with a Gaussian measure applied, thus allowing for convergence on a larger class of regions A. In particular, the proposed method would apply to unbounded regions, for which we cannot guarentee finite moments, convergence of the function $g(y)$ (let alone the explicit series expansion in terms of moments), or convergence of the Radon transform.

Here we will briefly discuss the potential challenges that this proposed method presents, which we address more carefully in sections 5 through 8. First and foremost instead of the standard moments (with respect to the Lebesgue measure), we are now given a Gaussian moment sequence — that is, moments with respect to the standard Gaussian measure on $\mathbb{R}^n$. On the other end, the reconstructed region may be recovered as normal by simply inverting the Gaussian weight.

From a theoretical standpoint, proving the validity of the proposed method presents a few challenges. Firstly, even a quick glance at the classical literature tells us that moment problems on bounded domains are substantially better behaved than their unbounded counterparts. Whereas all moment problems on a bounded interval are determinate, we now have the potential to run in to indeterminate problems. We show that we can guarentee determinacy in a Gaussian-weighted L^2 completion of the space of polynomials. This of course includes our primary use case: indicator functions.

Secondly, we must address the effect of the proposed modifications on the Radon transform. We will define the Gaussian Radon transform, which unlike the Radon transform exists for unbounded regions A.

The Stieltjes transform on unbounded domains is significantly more complicated than the original case. From the integral representation

$$g(y) = \int_{\mathbb{R}^n} \frac{f(x)}{1 + \langle x, y \rangle} \, dx,$$

one can start to see a problem: If $f(x)$ has unbounded support $g(y)$ has the potential to not converge for any $y \in \mathbb{R}^n$. In particular $g(y)$ may not exist in a neighborhood of 0, meaning our series expansion and thus the explicit conection to moments, may be broken. We need now to consider g as a function on some non-real domain. Here we show Pade approximants can be computed from moments which still approximate g on, for example, a half space. The convergence of these approximants is tied to the question determinacy of the moment problem, which we resolve as described above. Thus we simply take sample points avoiding the problematic real space and the method is validated.

Chapter 1

Background

1.1 Basic Notation

First recall the conventional multi-index notation. Let $\mathbb{N}_0$ denote the nonnegative integers. A multi-index $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_n) \in \mathbb{N}_0^n$ is an n-tuple of nonnegative integers. The degree of a multi-index is $|\alpha| = \alpha_1 + \alpha_2 + \cdots + \alpha_n$. Multivariate exponentiation is defined as follows. For $x = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$,

$$x^\alpha = x_1^{\alpha_1} x_2^{\alpha_2} \cdots x_n^{\alpha_n}.$$

Here we will use the convention $0^0 = 1$ so that for example $(x, y, z)^{(0,1,0)} = y$. We will use $e_i \in \mathbb{R}^n$ to represent the canonical unit vector

$$e_i = (0, \ldots, 0, 1, 0, \ldots, 0)$$

which can also be interpreted as a multi-index so for example

$$(x, y, z)^{2e_1 + e_3} = x^2 z$$

The multinomial formula gives a convenient expansion for multinomial powers. Let $k \in \mathbb{N}_0$. Then

$$(b_1 + b_2 + \cdots + b_n)^k = \sum_{|\alpha| = k} \binom{k}{\alpha} b^\alpha$$

where the multinomial coefficients are defined

$$\binom{k}{\alpha} = \frac{k!}{\alpha_1! \alpha_2! \cdots \alpha_n!}.$$

Note that the multinomial expansion sums over all multi-indices α of degree k.

We denote the standard euclidean inner product

$$\langle x, y \rangle = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n$$

where $y = (y_1, y_2, \ldots, y_n) \in \mathbb{R}^n$ and the Euclidean norm

$$\|x\| = \langle x, x \rangle = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}$$

We may slightly abuse notation by using the inner product and norm in different dimensions, even in the same equation. This can be excused if one views each euclidean space $\mathbb{R}^n$ as embedded in the sequence space $\ell^2(\mathbb{R}^{\mathbb{N}})$ in which the inner product and norm are equivalent.

When discussing hyperplanes in $\mathbb{R}^n$ we index them by a unit normal vector, $\omega \in S^{n-1}$, and distance from origin $-\infty < p < \infty$, and we write for example the implicit hyperplane equation $\langle x, \omega \rangle = p$. Note that $\langle x, -\omega \rangle = -p$ represents the same hyperplane. It is perhaps more correct, in later defining the Radon and Gaussian Radon transforms, to identify these indexes and define the transforms over a projective space. We omit this discussion as it is not relevant within the scope of this work.

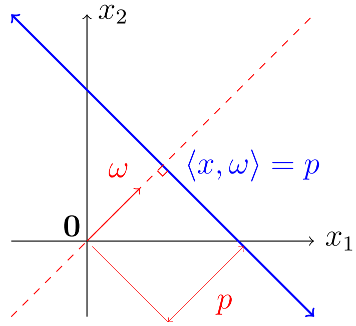

Figure 1.1.1: Hyperplane equation

fig:HypEq

Unless otherwise indicated all measures are Euclidean measures — that is the Borel measure associated with the standard Euclidean metric on a given region of integration. In particular the Euclidean measure on a hyperplane of $\mathbb{R}^n$ is equivalent to the Euclidean measure on $\mathbb{R}^{n-1}$. We choose for convenience to denote measures by their associated variable such as dx, dp, dz, and others. It should be stated that this notation, while uniform, is context dependent. For example in the integrals

$$\int_{\mathbb{R}^n} dx \qquad \text{and} \qquad \int_{\langle x,\omega\rangle = p} dx,$$

the measure dx is to be understood as the n-dimensional and $(n-1)$-dimensional Euclidean measure respectively.

1.2 Classical and multivariate moment problems

Let μ be a Borel measure on $\mathbb{R}$. For $k \in \mathbb{N}_0$, define the kth moment of μ as

$$c_k = \int_{-\infty}^{\infty} x^k \, d\mu(x)$$

provided the integral converges. The sequence $(c_k)_{k\in\mathbb{N}_0}$ is called a moment sequence or a moment problem, and the measure μ is called a solution to the moment problem. Loosely speaking a moment problem poses the question: Under given constraints (e.g. measures supported within a fixed subset of $\mathbb{R}^n$), to what extent can one determine a solution μ (if one exists) from a moment sequence?

The classical study of moment problems is divided into three cases depending on the support of μ. In each case there is a standard choice of support, from which general results are often derived by a change of variables.

Markov (or **Haussdorf**) moment problems within a bounded support.

$$\mathrm{supp}(\mu) \subseteq [0, 1]$$

Stieltjes moment problems within a one-sided infinite support.

$$\mathrm{supp}(\mu) \subseteq [0, \infty)$$

Hamburger moment problems within a two-sided infinite support.

$$\mathrm{supp}(\mu) \subseteq \mathbb{R}$$

In some special cases we will denote the moments c_k in a particular form:

(i) When μ is absolutely continuous with respect to the Lebesgue measure dx, moments can be defined by

$$c_k = \int f(x)x^k dx$$

where $f(x)$ is the density of μ with respect to dx.

(ii) When μ is the Lebesgue measure on a set A of finite measure, moments can be defined by

$$c_k = \int_A x^k dx.$$

We often use these representations interchangeably. For example, we may say a function f is determinate, or a set A is indeterminate, if the corresponding measure μ has that property.

There are two natural questions one can ask about a moment problem:

Solvability: Does a solution μ exist possessing the given moments?

Determinacy: Is a solution unique? If not, what can be said about the set of solutions?

For our purposes we will assume existence. In terms of the shape reconstruction method of Chapter 3 it is taken as a given that a solution exists, but that we have access only to the moments. Furthermore in practical application we encounter a followup question to determinacy: How can we reconstruct the solution?

In the classical cases (Markov, Stieltjes, Hamburger) these questions have been more or less resolved. Precise conditions for solvable and determinate moment problems have been found, and the nature of solution sets to indeterminate moment problems are well understood. For instance, all solvable Markov moment problems are unique. This result follows from the Weierstrass approximation theorem:

Proposition 1.2.1. All Markov moment problems are determinate. If μ, ν are Borel measures on the unit interval $[0,1]$ with equal moments,

$$\int_0^1 x^k d\mu = \int_0^1 x^k d\nu, \qquad k \in \mathbb{N}_0,$$

then $\mu = \nu$.

Example 1.2.2. An example of an indeterminate Stieltjes moment problem?

However, in the case of multivariate moment problems remain largely open. Let μ be a Borel measure now on $\mathbb{R}^n$. For $\alpha \in \mathbb{N}_0^n$. Define the αth multivariate moment of f as

$$c_\alpha = \int_{\mathbb{R}^n} x^\alpha \; d\mu(x).$$

Example 1.2.3. Let $A = [0,1]^n \subseteq \mathbb{R}^n$ be the unit cube and μ the Lebesgue measure on A. Then μ has multivariate moments

$$\begin{aligned}\int_{\mathbb{R}^n} x^\alpha \; d\mu(x) &= \int_{[0,1]^n} x_1^{\alpha_1} x_2^{\alpha_2} \cdots x_n^{\alpha_n} dx \\ &= \int_0^1 x_1^{\alpha_1} \; dx_1 \int_0^1 x_2^{\alpha_2} \; dx_2 \cdots \int_0^1 x_n^{\alpha_n} \; dx_n \\ &= \frac{1}{(\alpha_1+1)(\alpha_2+1)\cdots(\alpha_n+1)}\end{aligned}$$

Similarly the Lebesgue measure μ on any rectanglular region $A = \prod_{i=1}^n [a_i, b_i]$ has moments

$$\int_A x^\alpha \; d\mu(x) = \prod_{i=1}^n \int_{a_i}^{b_i} x_i^{\alpha_i} \; dx_i = \prod_{i=1}^n \frac{b_i^{\alpha_i+1} - a_i^{\alpha_i+1}}{\alpha_i + 1}$$

The moment sequence $(c_\alpha)_{\alpha \in \mathbb{N}_0^n}$ is now multi-indexed, but the notions of solvability and determinacy are analogous: It is solvable provided there exists a Borel measure with these moments, and it is determinate provided such a measure is unique. Here, we divide multivariate moment problems into two cases:

Bounded multivariate moment problems within a bounded support.

$$\mathrm{supp}(\mu) \subseteq [0,1]^n$$

Unbounded multivariate moment problems within an unbounded support.

$$\mathrm{supp}(\mu) \subseteq \mathbb{R}^n$$

While conditions for the solvability or determinacy of multivariate moment problems have been found in certain cases [**CITATION NEEDED**], they are nowhere near as comprehensive as those of classical moment problems. One straightforward, albeit limited, approach is to reduce and n-variate moment problem to a set of n univariate moment problems [1]. For example, we employ this strategy in our proofs of propositions [determinacy for bounded domain and gaussian measures].

As in the classical case of Markov moment problems, determinacy is a consequence of Stone-Weierstrass (n-variate polynomials):

Proposition 1.2.4. If μ is a Borel measure on the unit cube $A = [0,1]^n$, with moments

$$c_\alpha = \int_A x^\alpha d\mu, \qquad \alpha \in \mathbb{N}_0^n$$

then the moment problem $(c_\alpha)_{\alpha \in \mathbb{N}_0^n}$ is determinate.

We present a simple alternate proof in section 1.4 via the Radon transform.

Example 1.2.5. An example of an indeterminate, unbounded, multivariate moment problem?

1.3 Gaussian measures and the Hermite polynomials

We review the definitions of univariate and multivariate Gaussian measures and Hermite polynomials, and discuss some basic properties. An extensive investigation of Gaussian measures can be found in [Bogachev] [**CITATION NEEDED**]. The univariate Hermite polynomials are well known, but there are multiple multivariate generalizations. We use a simple product definition.

First we review some basic properties of the univariate Gaussian measure.

Definition 1.3.1. Let $w(p) : \mathbb{R} \to \mathbb{R}$ be the standard Gaussian density on $\mathbb{R}$,

$$w(p) = \frac{1}{\sqrt{2\pi}} e^{\frac{-p^2}{2}}$$

and $w_n(x) : \mathbb{R}^n \to \mathbb{R}$ the standard Gaussian density on $\mathbb{R}^n$,

$$w_n(x) = (2\pi)^{-\frac{n}{2}} e^{\frac{-\|x\|^2}{2}}.$$

Note that $w = w_1$.

A couple of formulas: First the Gaussian integral,

$$\int_{-\infty}^{\infty} w(p)\ dp = 1.$$

This can be proven in a number of clever ways.

Proof. We would like to show that

$$\int_{-\infty}^{\infty} e^{-\frac{p^2}{2}} dp = \sqrt{2\pi}$$

[**PROOF NEEDED**] □

Second: The derivative

$$w'(p) = -pw(p).$$

Third: The Gaussian function dominates polynomials in the sense that

$$\lim_{p \to -\infty} P(p)w(p) = 0 = \lim_{p \to \infty} P(p)w(p)$$

for any polynomial P.

Definition 1.3.2. The *Gaussian moments* of a function $f : \mathbb{R}^n \to \mathbb{R}$ are

$$c_\alpha^G = \int_{\mathbb{R}^n} f(x)x^\alpha w(x)\ dx$$

for $\alpha \in \mathbb{N}_0^n$.

Example 1.3.3. Let $f : \mathbb{R} \to \mathbb{R}$ be the constant function $f(p) = 1$. The Gaussian moments of f, i.e. the moments of $w(p)$, are

$$c_k^G = \begin{cases} (k-1)!!, & k \text{ even} \\ 0, & k \text{ odd} \end{cases}$$

We have already seen $c_0 = 1$. Furthermore it is not hard to see that for odd k,

$$c_k^G = \int_{-\infty}^{\infty} p^k w(p)\ dp = 0$$

since $p^k w(p)$ is an odd function.

Now for $\ell = 1, 2, \ldots$, we derive a recurrence formula for $c_{2\ell}$ by integrating by parts

$$\begin{aligned}
c^G_{2\ell} &= \int_{-\infty}^{\infty} p^{2\ell} w(p)\ dp \\
&= -\int_{-\infty}^{\infty} p^{2\ell-1} \left(-pw(p)\right)\ dp \\
&= -\left[p^{2\ell-1} w(p)\right]_{-\infty}^{\infty} + (2\ell-1)\int_{-\infty}^{\infty} p^{2\ell-2} w(p)\ dp \\
&= (2\ell-1) c^G_{2\ell-2}.
\end{aligned}$$

Thus by induction,

$$c^G_{2\ell} = (2\ell-1)(2\ell-3)\cdots c_0 = (2\ell-1)!!.$$

As an alternate derivation, for readers familiar with the gamma function $\Gamma(z) = \int_0^\infty t^{z-1} e^{-t}$ we present the following. Since $p^{2\ell} w(p)$ is even,

$$\int_{-\infty}^{\infty} p^{2\ell} e^{-\frac{p^2}{2}} dp = 2\int_0^\infty p^{2\ell} e^{-\frac{p^2}{2}} dp$$

Now substituting $t = \frac{p^2}{2}$, whereby $dt = p\ dp$,

$$\begin{aligned}
2\int_0^\infty p^{2\ell} e^{-\frac{p^2}{2}} dp &= 2^{\ell+\frac{1}{2}} \int_0^\infty t^{\ell-\frac{1}{2}} e^{-t} dt \\
&= 2^{\ell+\frac{1}{2}} \Gamma\left(\ell + \frac{1}{2}\right)
\end{aligned}$$

so that

$$c^G_{2\ell} = \frac{2^\ell}{\sqrt{\pi}} \Gamma\left(\ell + \frac{1}{2}\right) = (2\ell-1)!!$$

A relationship between the univariate and multivariate Gaussian densities can be seen as follows: Let $x, y \in \mathbb{R}^n$, $n \geq 2$. Suppose $y = p\omega$ where $\omega \in S^{n-1}$ and $p \in \mathbb{R}$, so that $p\omega$ is the orthogonal projection of x onto the span of y. Then the Pythagorean relation,

$$\|x\|^2 = \|x - p\omega\|^2 + \|p\omega\|^2 = \|x - p\omega\|^2 + p^2$$

implies that

$$\begin{aligned}
w_n(x) &= (2\pi)^{-\frac{n}{2}} e^{\frac{-\|x\|^2}{2}} \\
&= (2\pi)^{-\frac{n-1}{2}} e^{\frac{-\|x-p\omega\|^2}{2}} (2\pi)^{-\frac{1}{2}} e^{\frac{-p^2}{2}} \\
&= w_{n-1}(x - p\omega) w(p)
\end{aligned}$$

The equation

$$w_n(x) = w_{n-1}(x - p\omega)w(p) \tag{1.3.1}$$

eq:gaussPyt

is in some respect the defining property of the Gaussian measure described below. Indeed if $x = (x_1, x_2, \ldots, x_n)$, by repeated application of (eq:gaussPythag 1.3.1) one can write the decomposition

$$w_n(x) = \prod_{k=1}^{n} w(x_k).$$

Thus the $w_n(x)$ is the product of n copies of w. The standard Gaussian measure γ^n is thus a measure whose cardinal projections are standard Gaussian measures.

Definition 1.3.4. Let γ^n be the Borel measure on $\mathbb{R}^n$ with density w_n,

$$\int_{\mathbb{R}^n} f(x)d\gamma^n = \int_{\mathbb{R}^n} f(x)w_n(x)\ dx.$$

We call γ^n the *standard Gaussian measure*. More generally, for mean $a \in \mathbb{R}^n$, and variance $\sigma > 0$, the following γ^n_{a,σ^2} are *Gaussian measures*,

$$\begin{aligned}\int_{\mathbb{R}^n} f(x)d\gamma^n_{a,\sigma^2} &= \frac{(2\pi)^{-\frac{n}{2}}}{\sigma}\int_{\mathbb{R}^n} f(x)e^{-\frac{\|x-a\|^2}{2\sigma^2}}\ dx \\ &= \int_{\mathbb{R}^n} f(\sigma x + a)d\gamma^n\end{aligned}$$

Example 1.3.5. Let $f : \mathbb{R}^n \to \mathbb{R}$ be the constant function $f(x) = 1$. The Gaussian moments of f are the multivariate moments of $w_n(x)$, and can be written as

$$\begin{aligned}c^G_\alpha &= \int_{\mathbb{R}^n} x^\alpha w_n(x)\ dx \\ &= \prod_{i=1}^{n}\int_{-\infty}^{\infty} x_i^{\alpha_i} w(x_i)\ dx \\ &= \prod_{i=1}^{n} c^G_{\alpha_i}\end{aligned}$$

where $\{c_k\}_{k\in\mathbb{N}}$ are the univariate moments of $w(p)$. Thus $c^G_\alpha = 0$ if any α_i is odd. Otherwise, if every α_i is even then

$$c^G_\alpha = \prod_{i=1}^{n}(\alpha_i - 1)!!$$

Definition 1.3.6. The **Hermite polynomials** can be defined by the Rodrigues formula

$$H_k(p) = (-1)^k\frac{w^{(k)}(p)}{w(p)}$$

We can derive a recurrence relation for H_k with some basic differential calculus. First multiplying both sides by $(-1)^k w(p)$, we get

$$w^{(k)}(p) = (-1)^k w(p) H_k(p)$$

Now we differentiate both sides

$$w^{(k+1)}(p) = (-1)^k(-pw(p)H_k(p) + w(p)H_k'(p)) = (-1)^{k+1} w(p)(pH_k(p) - H_k'(p))$$

so that, dividing both sides by $(-1)^{k+1} w(p)$, we have

$$H_{k+1}(p) = pH_k(p) - H_k'(p)$$

Since $H_0(p)$ is trivially the constant 1 we compute the first handful of polynomials:

$$\begin{aligned}
H_0(p) &= 1 \\
H_1(p) &= p \\
H_2(p) &= p^2 - 1 \\
H_3(p) &= p^3 - 3p \\
H_4(p) &= p^4 - 6p^2 + 3 \\
H_5(p) &= p^5 - 10p^3 + 15p
\end{aligned}$$

Note that each H_k is a monic polynomial of degree k, and H_k is an even (odd resp.) function if k is even (odd resp.). The Hermite polynomials are mutually orthogonal in the sense that for nonnegative integers $\ell < k$,

$$\int_{-\infty}^{\infty} H_\ell(p) H_k(p) w(p)\ dp = 0 \tag{1.3.2}$$

eq:HOrth

To see this, substitute the Rodrigues formula for H_k,

$$(-1)^k \int_{-\infty}^{\infty} H_\ell(p) w^{(k)}(p)\ dp$$

and integrate by parts

$$\begin{aligned}
(-1)^k \left(\left[H_k(p) w^{(k-1)}(p)\right]_{-\infty}^{\infty} - \int_{-\infty}^{\infty} H_\ell'(p) w^{(k-1)}(p)\ dp \right) \\
= (-1)^{k-1} \int_{-\infty}^{\infty} H_\ell'(p) w^{(k-1)}(p)
\end{aligned}$$

where the boundary term $\left[H_k(p) w^{(k-1)}(p)\right]_{-\infty}^{\infty}$ vanishes because of $w(p)$'s rapid decay at $\pm\infty$. Repeat this process until H_ℓ is annihilated.

The Hermite polynomials form a complete orthogonal basis for $L^2(\mathbb{R}, w)$, the space of function $f : \mathbb{R} \to \mathbb{R}$ for which $\int_{-\infty}^{\infty} f(p)^2 w(p)\ dp < \infty$. This is a Hilbert space with inner product $\langle f, g\rangle = \int_{-\infty}^{\infty} |f(p)g(p)| w(p)\ dp$.

Proposition 1.3.7. The linear span of the Hermite polynomials $H_0(p), H_1(p), \ldots$

is dense in $L^2(\mathbb{R}, w)$.

Corollary 1.3.8. For any $f \in L^2(\mathbb{R})$ the moment problem given by

$$c_n = \int_{-\infty}^{\infty} f(p)w(p)dp$$

is determinate.

In the section 1.5 we will make use of a different definition for the Hermite polynomials, as Taylor series coefficients for an exponential generating function.

Lemma 1.3.9. The Hermite polynomials have the **generating function**,

$$\phi(p) = e^{pt-\frac{t^2}{2}} = \sum_{k=0}^{\infty} \frac{H_k(p)}{k!} t^k.$$

which converges absolutely for all $p, t \in \mathbb{R}$.

Proof. Note that

$$e^{pt-\frac{t^2}{2}} = \frac{e^{-\frac{(p-t)^2}{2}}}{e^{-\frac{p^2}{2}}} = \frac{w(p-t)}{w(p)}$$

so that the Taylor coefficients are precisely

$$\left[\frac{d^k}{dt^k}\frac{w(p-t)}{w(p)}\right]_{t=0} = (-1)^k \frac{w^{(k)}(p)}{w(p)} = H_k(p).$$

□

We can now define a straight forward multivariate analogue or the Hermite polynomials

Definition 1.3.10. For $\alpha \in \mathbb{N}_0^n$ we define the **multivariate Hermite polynomial** H_α by

$$H_\alpha(x) = H_{\alpha_1}(x_1) \cdots H_{\alpha_n}(x_n), \qquad x = (x_1, \dots x_n) \in \mathbb{R}^n$$

The multivariate Hermite polynomials have many useful properties corresponding to those of the classical polynomials. In particular we have the Rodrigues formula

$$H_\alpha(x) = (-1)^{|\alpha|} \frac{\partial^\alpha w_n(x)}{w_n(x)}$$

where $\partial^\alpha = \frac{\partial^{\alpha_1+\cdots+\alpha_n}}{\partial x_1^{\alpha_1}\cdots\partial x_n^{\alpha_n}}$. Just like the classical case, this formula can be used to define H_α as multivariate Taylor coefficients for the generating function:

$$\phi(x) = e^{\langle x,y\rangle - \frac{\|y\|^2}{2}} = \frac{w_n(x-y)}{w_n(x)} = \sum_{\alpha \in \mathbb{N}_0^n} \frac{H_\alpha(x)}{\alpha!} y^\alpha$$

which we note is absolutely convergent for all $x, y \in \mathbb{R}^n$. Here we use the same symbol ϕ for the univariate and multivariate generating functions at the risk of confusion. It should be clear from context.

1.4 The Radon and Gaussian Radon transforms

Let f be a function on the n-dimensional Euclidean space $\mathbb{R}^n$. We imagine taking "slices" of f by restricting it to a $(n-1)$-dimensional hyperplane Λ. These hyperplanes form the domain of the Radon Transform. More precisely, the Radon Transform associates each slice with a corresponding integral

$$R_f(\Lambda) = \int_\Lambda f(x)\ dx,$$

which can be thought of as a $(n-1)$-dimensional measurement of f. To be more precise we parametrize the collection of hyperplanes Λ by normal vector, $\omega \in S^{n-1}$, and (signed) distance from the origin $-\infty < p < \infty$. Indeed any hyperplane can be described in the form $\Lambda = \{x \in \mathbb{R}^n : \langle x, \omega\rangle = p\}$. As mentioned, there is a slight inconsistency in these definitions where a hyperplane Λ can be indexed by both $\langle \omega, x\rangle = p$ and $\langle -\omega, x\rangle = -p$. This difference is inconsequential for our purposes so we choose the latter definition for clarity.

The Radon transform [?Helg65] (RT) is a thing that I will discuss the history of, with references, in this paragraph. The transform gets its name from Johann Radon, whose first defined the transform in the form below in 1917, although a similar transform was introduced by Paul Funk in 1911 [**CITATION NEEDED**]. It is interesting to note that the primary application of the RT in medical imaging (CT scans) was not invented for another half century. I can only hope that in 2076 my dissertation will serve as an absorbent coffee coaster for a sleep deprived student.

Definition 1.4.1. Let f be a nonnegative measurable function on $\mathbb{R}^n$. The **Radon transform** $R_f : S^{n-1} \times \mathbb{R} \to \mathbb{R}$ of f is a function which, given a unit vector $\omega \in \mathbb{R}^n$ and $p \in \mathbb{R}$, is defined as

$$R_f(\omega, p) = \int\limits_{\langle x,\omega\rangle = p} f(x)\ dx,$$

provided the integral converges.

Example 1.4.2. For computation it is convenient to write the Radon Transform with an explicit isometric parameterization $x(t)$ of the hyperplane $\langle x, \omega\rangle = p$. In particular we note that there exists a map $x : \mathbb{R}^{n-1} \to \mathbb{R}^n$ such that $x(0) = p\omega$ and

$$R_f(\omega, p) = \int_{\mathbb{R}^{n-1}} f(x(t))dt.$$

For reference let's specify an hyperplane parameterization for the $n = 2$ case. In $\mathbb{R}^2$ often we identify ω with the angle $0 \leq \theta < 2\pi$ such that $\omega = (\cos\theta, \sin\theta)$. We define $x(t)$ by

$$x(t) = (t\sin\theta + p\cos\theta, -t\cos\theta + p\sin\theta), \qquad -\infty < t < \infty.$$

It is not difficult to show that $x(0) = p\omega$, $\langle x(t), \omega\rangle = p$, and most importantly

$$R_f(\omega, p) = \int_{-\infty}^{\infty} f(t\sin\theta + p\cos\theta, -t\cos\theta + p\sin\theta)\; dt$$

Example 1.4.3. Let $B(r) = \{x : |x| \leq r\} \subseteq \mathbb{R}^n$ be the ball of radius r.

$$R_{B(r)}(\omega, p) = \begin{cases} V_{n-1}\left(\sqrt{r^2 - p^2}\right), & |p| \leq r \\ 0 & |p| > r \end{cases}$$

where $V_n(r)$ is the volume of a hypersphere of radius r,

$$V_n(r) = \frac{\pi^{n/2}}{\Gamma(\frac{n}{2} + 1)} r^n.$$

Note that as a bounded domain, we can guarantee $B(r)$ is integrable on all hyperplanes. In fact the RT of any subset of $B(r)$ is sharply bounded by $R_{B(r)} \leq V_{n-1}(r)$. Also note that $B(r)$ is rotation invariant, hence $R_{B(r)}(\omega, p)$ is independent of ω.

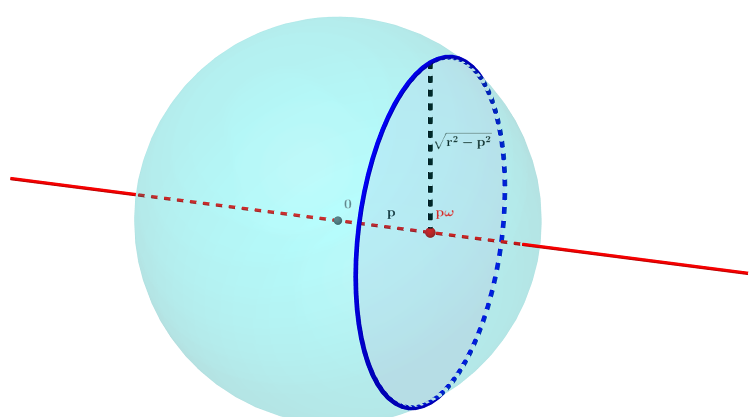

Figure 1.4.1: The RT of a ball

fig:RTBall

We can use this formula to determine the RT of an annulus. Let $A(r_1, r_2) = \{x \in$

$\mathbb{R}^n : r_1 \leq \|x\| \leq r_2\}$. Then

$$\begin{aligned} R_{A(r_1,r_2)}(\omega, p) &= R_{B(r_2)}(\omega, p) - R_{B(r_1)}(\omega, p) \\ &= \begin{cases} V_{n-1}(\sqrt{r_2^2 - p^2}) - V_{n-1}(\sqrt{r_1^2 - p^2}), & |p| < r_1 \\ V_{n-1}(\sqrt{r_2^2 - p^2}), & r_1 \leq |p| \leq r_2 \\ 0, & |p| \geq r_2 \end{cases} \end{aligned}$$

Now consider an example of an unbounded region.

Example 1.4.4. Let $S = \{(x, y) : |y| \leq \frac{1}{2}\} \subseteq \mathbb{R}^2$ be a strip centered on the x-axis with width 1. Clearly if $\theta = \pi/2$ or $3\pi/2$ then

$$R_S(\theta, p) = \begin{cases} \infty, & |p| \leq \frac{1}{2} \\ 0, & |p| > \frac{1}{2} \end{cases}.$$

Otherwise,

$$R_S(\theta, p) = \sec\theta$$

Note because S is unbounded, that R_S is not only unbounded, but even divergent in some cases.

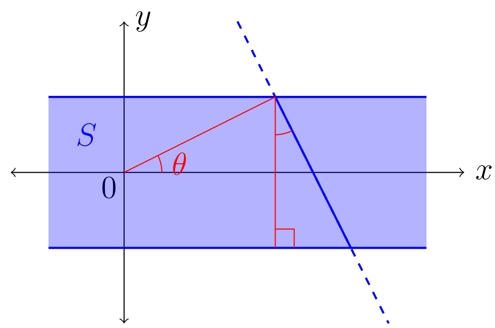

Figure 1.4.2: RT of a strip fig:strip

Our main addition to previous work will be the use of a modified RT, the Gaussian Radon transform (GRT). This transform is very similar to the RT, but the inclusion of a Gaussian density $w_{n-1}(x)$ in the integral allows for convergence on a larger class of functions f. This includes for example unbounded regions. In a broader context the Gaussian Radon transform also has the advantages of generalizing to infinite dimensional Hilbert spaces [?Seng14] (on which the Lebesgue measure is not defined), as well as having a natural probabilistic interpretation.

Definition 1.4.5. The *Gaussian Radon transform* $GR_f : S^{n-1} \times \mathbb{R} \to \mathbb{R}$ of f is defined similarly to the RT. Given $\omega \in S^{n-1}$ and $-\infty < p < \infty$, the GRT is

$$GR_f(\omega, p) = \int\limits_{\langle x, \omega \rangle = p} f(x) w_{n-1}(x - p\omega)\, dx.$$

provided the integral converges. Note the Gaussian density

$$w_{n-1}(x - p\omega) = (2\pi)^{-(n-1)/2} e^{-\|x-p\omega\|^2/2}$$

is centered on the point of the hyperplane closest to the origin.

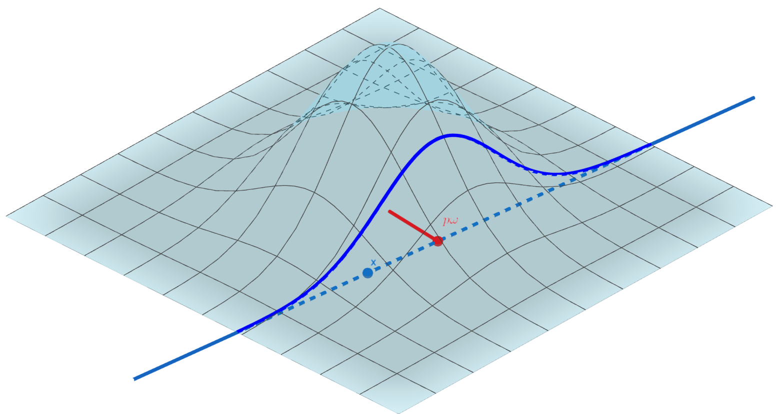

Figure 1.4.3: The GRT fig:GRT

Remark 1.4.6. It may be helpful to understand the GRT as a simple modification of the RT with respect to a Gaussian measure on $\mathbb{R}^n$. From the relation

$$\int_{\langle x,\omega\rangle = p} f(x) w_n(x)\ dx = \int_{\langle x,\omega\rangle = p} f(x) w_{n-1}(x)\ dx\ w(p),$$

we can express the GRT of f in terms of the *RT* of the function $g(x) = f(x)w_n(x)$:

$$R_g(\omega, p) = GR_f(\omega, p) w(p), \qquad g(x) := f(x) w_n(x). \tag{1.4.1}$$ eq:GRTPytha

The relation above provides decent intuition for the GRT, and is also a useful tool proving some basic properties of the transform. Expand on intuition (maybe with example.)

Example 1.4.7. If $x(t) : \mathbb{R}^{n-1} \to \mathbb{R}^n$ is a parametrization of $\langle x, \omega\rangle = p$ as described above, then

$$GR_f(\omega, p) = \int_{\mathbb{R}^{n-1}} f(x(t)) w_{n-1}(t) dt$$

In particular for $f : \mathbb{R}^2 \to \mathbb{R}$

$$GR_f(\omega, p) = \int_{-\infty}^{\infty} f(t\sin\theta + p\cos\theta, -t\cos\theta + p\sin\theta) w(t)\ dt$$

where $\omega = (\cos\theta, \sin\theta)$.

Example 1.4.8. The Gaussian Radon transform is bounded for any measurable region $A \subseteq \mathbb{R}^n$. Indeed

$$GR_A(\omega, p) \le GR_1(\omega, p) = \int_{\mathbb{R}^{n-1}} w_{n-1}(t) dt w(p) = w(p)$$

Example 1.4.9. Consider the Strip $S \subseteq \mathbb{R}^2$ from [earlier example]. While the RT of S was divergent for $\theta = \pi/2$ or $3\pi/2$, the GRT converges. In particular,

$$GR_S(\theta, p) = \begin{cases} \frac{1}{\sqrt{2\pi}} e^{-\frac{p^2}{2}}, & |p| \le \frac{1}{2} \\ 0, & |p| > \frac{1}{2} \end{cases}$$

This is the density of a "truncated" normal distribution on $[-\frac{1}{2}, \frac{1}{2}]$. For other angles $GR_S(\theta, p)$ is likewise the density of a truncated normal distribution, this time with arbitrary endpoints $[a, b]$, given by

$$a = ?, \qquad b = a + \sec\theta$$

Now imagine sweeping a hyperplanar "slice" across $\mathbb{R}^n$. As a function of p, $R(\omega, p)$ can be seen as a projection of f onto the linear subspace spanned by ω. It is not surprising that integrating this projection over $-\infty < p < \infty$ we get the same result as the n-fold integral of f over $\mathbb{R}^n$.

$$\int_{-\infty}^{\infty} R(\omega, p)\ dp = \int_{\mathbb{R}^n} f(x)\ dx$$

The so called "slice theorem" further generalizes this observation:

Proposition 1.4.10 (Slice Theorem)**.** If $f \in L^1(\mathbb{R}^n)$ and $F \in L^\infty(\mathbb{R})$, then

$$\int_{-\infty}^{\infty} R_f(\omega, p) F(p) dp = \int_{\mathbb{R}^n} f(x) F(\langle x, \omega\rangle) dx, \tag{1.4.2}$$ eq:ST

Proof. This is a corollary of the Fubini-Tonelli theorem, which guarantees the slice formula as the integrals on either side converge absolutely. To see this note that there is a rigid transformation (Hilbert space isomorphism) taking this to an iterated integral over $\mathbb{R}^{n-1}$ and $\mathbb{R}$. Further, $F(\langle x, \omega\rangle) \in L^\infty(\mathbb{R}^n)$, so the proposition follows by Hölder's inequality

$$\int_{\mathbb{R}^n} |f(x) F(\langle x, \omega\rangle)|\ dx \le \int_{\mathbb{R}^n} |f(x)|\ dx \|F(\langle x, \omega\rangle)\|_\infty < \infty$$

□

Remark 1.4.11. The sufficient conditions for the slice theorem above can loosened significantly. We can take for example the straightforward Fubini condition

$$\int_{\mathbb{R}^n} |f(x) F(\langle x, \omega\rangle)| < \infty$$

which is necessary **[CITATION NEEDED]**, but not convenient. We may also use the condition f has bounded support and $F \in L^1_{loc}(\mathbb{R})$.

The conditions on f and F can be reduced somewhat. In general the convergence of either the left or right side above is sufficient. One

If $F(p) = e^{-ip}$ and $f(x)$ is such that $\int_{-\infty}^{\infty} R_f(\omega, p)dp < \infty$ then (1.4.2) becomes the well known Fourier slice theorem

$$\int_{-\infty}^{\infty} R_f(\omega, p)e^{-ip}\ dp = \int_{\mathbb{R}^n} f(x)e^{-i\langle x,\omega\rangle}\ dx,$$

which is often articulated as saying that the 1-dimensional Fourier transform of the Radon transform is the n-dimensional Fourier transform of f.

An early and natural question in the study of the RT is that of inversion. Radon himself derived the "Radon inversion formula" [?Rado17] [?Rado86], which is often proved via the above Fourier slice theorem. The groundbreaking inversion formula is the basis for what, in application, called "filtered back-propagation".

If one is interested in inverting the RT then a prerequisite concern is of course: Is the transform injective? The answer clearly depends on what space we draw the function f from. Radon [?Rado17] [?Rado86] provides a set of sufficient regularity conditions such that the RT is invertible. Other similar results followed [?????]. On the other hand counterexamples have been constructed by, for example, Lawrence Zalcman [?Zalc82], of continuous and nontrivial functions for which the RT is identically zero.

By way of the relation (1.4.1) we can prove an analogous slice theorem for the GRT. Mihai and Sengupta prove the general result if real, separable, infinite Hilbert spaces **[CITATION NEEDED]**.

Proposition 1.4.12 (Gaussian Slice Theorem)**.** Let $f : \mathbb{R}^n \to \mathbb{R}$ and $F : \mathbb{R} \to \mathbb{R}$ be measurable functions. If $f \in L^1(\mathbb{R}^n, w_n)$ and $F \in L^\infty(\mathbb{R})$ then

$$\int_{-\infty}^{\infty} GR_f(\omega, p)F(p)w(p)\ dp = \int_{\mathbb{R}^n} f(x)F(\langle x,\omega\rangle)w_n(x)\ dx. \qquad (1.4.3)$$

eq:GST

Remark 1.4.13. Again the more general Fubini condition

$$\int_{\mathbb{R}^n} |f(x)F(\langle x,\omega\rangle)w_n(x)|dx < \infty$$

may be used. Further, whatever conditions on f and F are sufficient for convergence in (1.4.2) are then sufficient conditions to be checked of $f(x)w_n(x)$ and $F(p)w(p)$. I'll need to check these conditions more carefully.

Proof. From (eq:GRTPythag 1.4.1)

$$\int_{-\infty}^{\infty} GR_f(\omega, p)F(p)w(p)\ dp = \int_{-\infty}^{\infty} R_g(\omega, p)F(p)\ dp$$

where $g(x) = f(x)e^{-\|x\|^2/2}$. Note that $\|g\|_1 = \|f\|_{1,w_n} < \infty$ and. Then applying the slice theorem:

$$\int_{-\infty}^{\infty} R_g(\omega, p)F(p)\ dp = \int_{\mathbb{R}^n} f(x)F(\langle x, \omega\rangle)w_n(x)\ dx,$$

completing the proof. $\square$

1.5 The Radon transform of measures

Taking from the context of classical moment problems we should define the notion of the Radon transform of a measure μ. Let μ be a Borel measure on $\mathbb{R}^n$ with finite moments c_α, $\alpha \in \mathbb{N}_0^n$. The projection $\pi_\omega : \mathbb{R}^n \to \mathbb{R}$ given by

$$\pi_\omega(x) = \langle x, \omega\rangle$$

is a Borel function, thus we may define the push-forward measure $\mu_\omega = \mu \circ \pi_\omega^{-1}$ which for Borel sets $A \subseteq \mathbb{R}$ is given by,

$$\mu_\omega(A) = \mu(\pi_\omega^{-1}(A)) = \mu(\{x : \langle x, \omega\rangle \in A\}).$$

We may call μ_ω the **marginal projection measure** of μ with direction vector ω.

Notice that if μ is absolutely continuous with density representation $\mu = f(x)dx$ then the slice theorem with $F(p)$ the characteristic function of A gives,

$$\mu(\pi_\omega^{-1}(A)) = \int_{\langle x,\omega\rangle \in A} f(x)dx = \int_A R_f(\omega, p)dp$$

so that in fact μ_ω is absolutely continuous with respect to the Lebesgue measure dp and it's density is precisely the Radon transform $R_f(\omega, p)$. In this sense the marginal projection μ_ω is a natural generalization of the RT, so we define

Definition 1.5.1. The Radon transform of a Borel measure μ on $\mathbb{R}$ is defined as the push-forward measure $\mu_\omega = \mu \circ \pi_\omega^{-1}$ on $\mathbb{R}$, where $\pi_\omega : \mathbb{R}^n \to \mathbb{R}$ is the projection

$$\pi_\omega(x) = \langle x, \omega\rangle.$$

We may denote this measure by $R_\mu^\omega = \mu_\omega$.

Remark 1.5.2. It seems that the slice theorem is analogous to — and perhaps generalized by — the change of variables formula for the push-forward measure

R^ω_μ:

$$\int_{-\infty}^{\infty} F(p)dR^\omega_\mu(p) = \int_{\mathbb{R}^n} F(\pi_\omega(x))d\mu$$

where $F : \mathbb{R} \to \mathbb{R}$ is integrable with respect to dR^ω_μ if and only if $F \circ \pi_\omega : \mathbb{R}^n \to \mathbb{R}$ is integrable with respect to μ.

Can we similarly define the GRT of a measure?

1.6 The Stieltjes transform

Let μ be a Borel measure with finite moments $c_k = \int_{-\infty}^{\infty} p^k d\mu$.

Definition 1.6.1. The complex function g defined by

$$g(z) = \int_{-\infty}^{\infty} \frac{d\mu}{1 + zp}$$

is called the Markov (Hamburger resp.) transform of μ when μ has bounded (unbounded resp.) support.

It is worth noting at the outset that $g(z)$ is defined at least for non-real z. To see this we note that if Im $z \neq 0$,

$$\begin{aligned} \frac{1}{|1 + zp|} &= \frac{|1/z|}{|\frac{1}{z} + p|} \\ &\leq \frac{|1/z|}{|\text{Im } \frac{1}{z}|} \\ &= \frac{|\overline{z}|/|z|^2}{|\text{Im } \overline{z}|/|z|^2} \\ &= \frac{|z|}{|\text{Im } z|} \end{aligned}$$

which gives the bound

$$|g(z)| \leq \int_{-\infty}^{\infty} \frac{|z|}{|\text{Im } z|} d\mu(p) = \frac{|z|}{|\text{Im } z|} c_0 \tag{1.6.1}$$

Since the integrand has a singularity only when $p = -1/z$ a Markov transform must converge in some neighborhood of zero. This cannot be said in general for a Hamburger transform.

Remark 1.6.2. The bound above has a useful geometric interpretation. For starters we note that,

$$\frac{|\text{Im } z|}{|z|} = |\sin(\arg z)|,$$

which can be seen by considering the right triangle formed by the points 0, z, and Re z. Thus this bound essentially depends on the angle of separation between z and the real line.

Proposition 1.6.3. The function $g(z)$ is analytic on the upper and lower half planes $\text{Im } z \neq 0$.

Proof. Let z, w be complex numbers with non-zero imaginary part. We have

$$\begin{aligned} g(w) - g(z) &= \int_{-\infty}^{\infty} \frac{1}{1+wp} - \frac{1}{1+zp} d\mu(p) \\ &= \int_{-\infty}^{\infty} \frac{(1+zp)-(1+wp)}{(1+wp)(1+zp)} d\mu(p) \\ &= (w-z)\int_{-\infty}^{\infty} \frac{-p}{(1+wp)(1+zp)} d\mu(p). \end{aligned}$$

Thus it seems

$$\lim_{w \to z} \frac{g(w)-g(z)}{w-z} = \int_{-\infty}^{\infty} \frac{-p}{(1+zp)^2} d\mu(p).$$

To be precise we may justify the interchange of limit and integral here by dominated convergence. Recall that it suffices to find an integrable function $h : \mathbb{R} \to \mathbb{C}$ such that, for any w in a neighborhood of z,

$$\left| \frac{-p}{(1+wp)(1+zp)} \right| \leq h(p) \qquad \text{for all } p \in \mathbb{R}.$$

It should be noted that our particular dominating function $h(p)$ is somewhat arbitrary, and not a sharp bound. By the previously discussed bound we have,

$$\left| \frac{-p}{(1+wp)(1+zp)} \right| \leq |p| \frac{|wz|}{|\text{Im } w \text{ Im } z|}.$$

Furthermore $|w|/|\text{Im } w|$ is clearly continuous on the half planes, so in some neighborhood of

$$|p| \frac{|wz|}{|\text{Im } w \text{ Im } z|} \leq |p| \left(\frac{|z|^2}{|\text{Im } z|^2} + \epsilon \right)$$

for an arbitrary $\epsilon > 0$. Finally, we note that the right hand dominating function above is integrable since μ has finite moments:

$$\int_{-\infty}^{\infty} |p| d\mu \leq \int_{-\infty}^{\infty} \frac{1}{2}(p^2+1)\ d\mu = \frac{c_2+c_0}{2} < \infty$$

$\square$

Remark 1.6.4. Since μ is a real measure, $g(z)$ commutes with conjugation,

$$g(\bar{z}) = \int_{\infty}^{\infty} \frac{d\mu}{1+\bar{z}p} = \overline{\int_{\infty}^{\infty} \frac{d\mu}{1+zp}} = \overline{g(z)}.$$

For this reason, although $g(z)$ is defined for all non-real z, we will only concern ourselves with its values on the upper half plane $\operatorname{Im} z > 0$ going forward. In chapter 3 we should restrict our sample points to the upper half plane, since samples in the lower half may be redundant.

When z is a real number $g(z)$ may not converge. In particular $g(z)$ does not converge when $z = -1/p$ for some p in the support of μ [**PROOF NEEDED**]. Thus if μ has bounded support g converges in a neighborhood of 0. On the other hand if μ has unbounded support $g(0)$ is not defined. However — as we will see — even in the Hamburger transform case, a formal series expansion at $z = 0$ holds valuable information. The results of this section are described in detail in section 5.6 of (Baker) [**CITATION NEEDED**]

By a geometric series expansion,

$$\frac{1}{1+zp} = \sum_{k=0}^{\infty} (-z)^k p^k$$

when $|zp| < 1$. This suggests the connection between $g(z)$ and the moment sequence $(c_k)_{k\in\mathbb{N}_0}$. Indeed, at least formally,

$$\begin{aligned}\int_{\infty}^{\infty} \frac{d\mu}{1+zp} &= \int_{\infty}^{\infty} \sum_{k=0}^{\infty} (-z)^k p^k \; d\mu \\ &= \sum_{k=0}^{\infty} (-z)^k \int_{\infty}^{\infty} p^k \; d\mu \\ &= \sum_{k=0}^{\infty} c_k (-z)^k.\end{aligned}$$

In the Markov case we have a positive radius of convergence. But even in the Hamburger case, there is a sense in which this asymptotic expansion holds "non-tangentially".

Non-tangential convergence at $z = 0$ is, as the name implies, convergence along curves not tangent to the real line at 0. More precisely, one defines a "wedge domain"

$$\Gamma_\delta = \{\delta \le \arg z \le \pi - \delta\}$$

in which curves approach 0 with an angle at least δ from the real line.

Definition 1.6.5. (Non-tangential convergence and asymptotic expansion) We

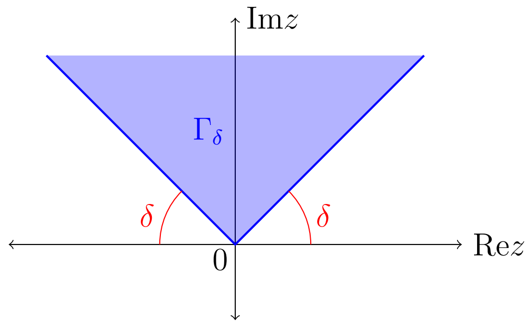

Figure 1.6.1: A wedge domain

fig:wedge

say $h(z) \to 0$ non-tangentially as $z \to 0$ if, for any $\delta > 0$, the limit

$$\lim_{z \to 0} h(z) = 0$$

holds for z in the Γ_δ. If a formal power series $\sum_{k=0}^{\infty} a_k z^k$ and a function $g(z)$ are such that for all $n \in \mathbb{N}_0$,

$$g(z) = \sum_{k=0}^{n-1} c_k (-z)^k + z^n h_n(z)$$

where the trailing terms $h_n(z)$ converge to 0 non-tangentially as $z \to 0$, we say that $\sum_{k=0}^{\infty} a_k z^k$ is an asymptotic expansion of $g(z)$. We use the notation

$$g(z) \simeq \sum_{k=0}^{\infty} a_k z^k$$

for asymptotic expansions.

While $g(z)$ does not generally converge at $z = 0$, in a non-tangential sense the aforementioned asymptotic expansion holds:

Proposition 1.6.6. (i) If μ has bounded support

$$g(z) = \sum_{k=0}^{\infty} c_k (-z)^k$$

in a neighboorhood of $z = 0$.

(ii) For all $n \in \mathbb{N}_0$

$$g(z) = \sum_{k=0}^{n-1} c_k (-z)^k + z^n h_n(z) \tag{1.6.2}$$

assexp

where h_n is such that $\lim_{z \to 0} h_n(z) = 0$ non-tangentially.

Proof. We first note that

$$\sum_{k=0}^{n} c_k(-z)^k = \sum_{k=0}^{n} \int_{-\infty}^{\infty} (-zp)^k d\mu = \int_{-\infty}^{\infty} \frac{1-(-zp)^{n+1}}{1+zp} d\mu$$

and so

$$h_n(z) = \frac{1}{z^n}\left\{ g(z) - \sum_{k=0}^{n} c_k(-z)^k \right\} = \int_{-\infty}^{\infty} \frac{zp^{n+1}}{1+zp} d\mu.$$

It remains to determine if and in what sense this trailing term $h_n(z)$ vanishes at the origin.

As expected in the case when μ has bounded support, $h_n(z)$ exists in a neighborhood of 0 and vanishes as $z \to 0$ in this neighborhood. Thus the asymptotic expansion assexp 1.6.2 is the Taylor expansion: The Markov transform is analytic at the origin.

When μ has unbounded support and $h_n(z)$ may not be well defined on the real line, so we must weaken our result to non-tangential convergence. If $z \in \Gamma_\delta$ then we have the following inequalities for any $p \in \mathbb{R}$,

$$|p-z| \geq |p| \sin\delta \quad \text{and} \quad |p-z| \geq |p| \sin\delta.$$

We will show that

$$\lim_{z\to 0} h_{2n}(z) = 0, \quad z \in \Gamma_\delta$$

from which the corresponding limit for odd orders will immediately follow from the relation

$$h_{2n-1}(z) = zh_{2n}(z) + c_{2n}z^{2n}.$$

Here for the sake of convenience we define $w = -\frac{1}{z}$, following a more classical approach. Note that $|w| = \frac{1}{|z|}$ and if z is in the wedge Γ_δ then so is w. Now

$$\begin{aligned} |h_{2n}(z)| &\leq \int_{-\infty}^{\infty} \frac{|p|^{2n+1}}{|p-w|} d\mu \\ &\leq \frac{|z|}{\sin\delta} \int_{|p|\leq A} |p|^{2n+1} d\mu + \frac{1}{\sin\delta} \int_{|p|\geq A} p^{2n} d\mu \\ &\leq \frac{2|z|A^{2n+2}}{\sin\delta} + \frac{1}{\sin\delta} \int_{|p|\geq A} p^{2n} d\mu \end{aligned}$$

Thus

$$\lim_{z\to 0} |h_{2n}(z)| \leq \frac{1}{\sin\delta} \int_{|p|\geq A} p^{2n} d\mu, \quad z \in \Gamma_\delta.$$

Since A is arbitrary and the integral $\int p^{2n} d\mu = c_{2n}$ is convergent we are done. $\square$

Remark 1.6.7. An interesting connection — which I don't know enough about to discuss yet — is beginning to be revealed here between moment problems and analytic functions on the upper half plane (in particular, their behavior near the real boundary). This leads us to the notion of interpolation problems [**CITATION NEEDED**].

Evidently the asymptotic series expansion $g(z) \simeq \sum_{k=0}^{\infty} c_k (-z)^k$ is generally formal, having zero radius of convergence except in the Markov case. However as we will see in the next section, rational functions constructed from this formal series exist which approximate $g(z)$ on the upper half plane $\{\mathrm{Im} z > 0\}$.

Finally we prove two propositions regarding the Hamburger transform of a RT projection. Let $\omega \in S^{n-1}$ be fixed and consider

$$g(z) = \int_{\infty}^{\infty} \frac{R(\omega, p)}{1 + zp} dp$$

for some function $f : \mathbb{R}^n \to \mathbb{R}$.

Definition 1.6.8. The multivariate Markov (Hamburger resp.) transform of a Borel measure μ on $\mathbb{R}^n$ of bounded (unbounded resp.) support is defined by

$$g(y) = \int_{\mathbb{R}^n} \frac{d\mu(x)}{1 + \langle x, y \rangle} \simeq \sum_{\alpha \in \mathbb{N}_0^n} c_\alpha (-y)^\alpha$$

when μ has finite moments $c_\alpha = \int_{\mathbb{R}^n} x^\alpha d\mu$.

By applying the slice theorem, with $F(p) = (1 + zp)^{-1}$ we see that the Hamburger transform of a projection can be represented by a similar multivariate integral of f over $\mathbb{R}^n$.

$$\int_{-\infty}^{\infty} \frac{R_f(\omega, p)}{1 + zp} dp = \int_{\mathbb{R}^n} \frac{f(x)}{1 + z\langle x, \omega \rangle} dx$$

Similarly the GRT slice theorem with $F(p) = (1 + zp)^{-1}$ is

$$\int_{-\infty}^{\infty} \frac{GR_f(\omega, p) w(p)}{1 + zp} dp = \int_{\mathbb{R}^n} \frac{f(x) w_n(x)}{1 + z\langle x, \omega \rangle} dx.$$

1.7 Padé approximants

In this section we recall the necessary definitions and results. Then we will prove some new results to be applied to the Gaussian Radon transform later.

Definition 1.7.1 (Classical definition)**.** The Padé approximant to a (possibly formal) power series

$$R^{[L/M]}(z) \simeq \sum_{k=0}^{\infty} c_k z^k$$

is a rational function with numerator (denominator resp.) degree at most L (M resp.), with series equal to $\sum_{k=0}^{N} c_k z^k + O(z^{N+1})$ up to as high an order N as possible.

Let

$$R^{[L/M]}(z) = \frac{P^{[L/M]}(z)}{Q^{[L/M]}(z)} = \frac{a_L z^L + \cdots + a_1 z + a_0}{b_M z^M + \cdots + b_1 z + b_0}.$$

Notice that in general there is a negligible constant common factor between the numerator and denominator, so that with the remaining $L+M+1$ free parameters, we expect an order of accuracy of up to $L+M+1$ constraints, $c_0, c_1, \ldots, c_{L+M}$. Thus we define the $[L/M]$ Padé approximant by the condition,

$$\frac{P^{[L/M]}(z)}{Q^{[L/M]}(z)} = \sum_{k=0}^{L+M} c_k z^k + O\left(z^{L+M+1}\right). \tag{1.7.1}$$

eq:BPade

Multiplying by $Q^{[L/M]}(z)$ gives a necessary condition, with a subtle but important difference from (1.7.1)

$$P^{[L/M]}(z) = Q^{[L/M]}(z) \left(\sum_{k=0}^{L+M} c_k z^k \right) + O\left(z^{L+M+1}\right). \tag{1.7.2}$$

eq:CPade

In the classical theory of Padé approximants (1.7.2) was often taken as a definition. It is always possible to find polynomials of the required degree satisfying this second condition, however they do not necessarily attain the degree of accuracy required by the first. We follow Baker, defining $R^{[L/M]}(z)$ by (1.7.1), provided such a rational function exists.

Definition 1.7.2 (Baker's definition)**.** The Padé approximant $R^{[L/M]}$ to a (possibly formal) power series is a the unique rational function with numerator (denominator resp.) degree at most L (M resp.) satisfying the condition (1.7.1)

$$\frac{P^{[L/M]}(z)}{Q^{[L/M]}(z)} = \sum_{k=0}^{L+M} c_k z^k + O\left(z^{L+M+1}\right).$$

If no such rational function exists we say the Padé approximant does not exist.

Note that the two definitions are equivalent if $b_0 = Q^{L/M}(0) \neq 0$ [**PROOF NEEDED**]. Hence a sufficient condition for the existence of the Padé approximant by Baker's definition is that $b_0 = Q^{L/M}(0) \neq 0$.

Equating coefficients of z^k, $k = 0, 1, \dots, L+M$ in (1.7.2) gives two linear systems

$$
\begin{aligned}
a_0 &= b_0 c_0 \\
a_1 &= b_1 c_0 + b_0 c_1 \\
&\vdots \\
a_L &= b_L c_0 + b_{L-1} c_1 + \cdots + b_0 c_L
\end{aligned}
$$

and,

$$
\begin{aligned}
0 &= b_M c_{L-M+1} + b_{M-1} c_{L-M+2} + \cdots + b_0 c_{L+1} \\
0 &= b_M c_{L-M+2} + b_{M-1} c_{L-M+3} + \cdots + b_0 c_{L+2} \\
&\vdots \\
0 &= b_M c_L + b_{M-1} c_{L+1} + \cdots + b_0 c_{L+M}
\end{aligned}
$$

where for convenience we set $c_k = 0$ for $k < 0$. The first system shows the numerator $P^{[L/M]}$ is determined by the denominator $Q^{[L/M]}$ (explicitly computing the each coefficient $a_1, \dots, a_L$ from $b_1, \dots b_L$). The second linear system is homogeneous, with M equations in $M+1$ unknowns. Thus we are guaranteed a nontrivial solution, and hence a Padé approximant by the classical definition. A clever modification of the latter system gives us a determinantal formula for $Q^{[L/M]}$. Augmented with the desired definition,

$$
Q^{[L/M]}(z) = b_M z^M + b_{M-1} z^{M-1} + \cdots + b_0
$$

the system (now with dimension $M+1 \times M+1$) can be written in vector form,

$$
\begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ Q^{[L/M]}(z) \end{pmatrix} = \begin{pmatrix} c_{L-M+1} & c_{L-M+2} & \cdots & c_{L+1} \\ c_{L-M+2} & c_{L-M+3} & \cdots & c_{L+2} \\ \vdots & \vdots & \ddots & \vdots \\ c_L & c_{L+1} & \cdots & c_{L+M} \\ z^M & z^{M-1} & \cdots & 1 \end{pmatrix} \begin{pmatrix} b_M \\ b_{M-1} \\ \vdots \\ b_1 \\ b_0 \end{pmatrix}
$$

Solving for b_0 by Cramer's Rule we get,

$$
b_0 \begin{vmatrix} c_{L-M+1} & c_{L-M+2} & \cdots & c_{L+1} \\ c_{L-M+2} & c_{L-M+3} & \cdots & c_{L+2} \\ \vdots & \vdots & \ddots & \vdots \\ c_L & c_{L+1} & \cdots & c_{L+M} \\ z^M & z^{M-1} & \cdots & 1 \end{vmatrix} = \begin{vmatrix} c_{L-M+1} & c_{L-M+2} & \cdots & c_L \\ c_{L-M+2} & c_{L-M+3} & \cdots & c_{L+1} \\ \vdots & \vdots & \ddots & \vdots \\ c_L & c_{L+1} & \cdots & c_{L+M-1} \end{vmatrix} Q^{[L/M]}(z).
$$

The LHS is a polynomial by Laplace expansion, and is nonzero if the RHS minor has nonzero determinant. Thus if the determinant

$$\begin{vmatrix} c_{L-M+1} \cdots & c_L \\ \vdots \quad \ddots & \vdots \\ c_L \quad \cdots & c_{L+M-1} \end{vmatrix}$$

is nonzero, then up to the aforementioned constant factor,

$$Q^{[L/M]}(z) = \begin{vmatrix} c_{L-M+1} \cdots & c_{L+1} \\ \vdots \quad \ddots & \vdots \\ c_L \quad \cdots & c_{L+M} \\ z^M \quad \cdots & 1 \end{vmatrix} \tag{1.7.3}$$ eq:Qdet

By convention we normalize $R^{[L/M]}(z)$ so that $b_0 = Q^{[L/M]}(0) = 1$. In terms of the determinental formula above, we can write $Q^{[L/M]}(0)$ as

$$Q^{[L/M]}(0) = \begin{vmatrix} c_{L-M+1} \cdots & c_L & c_{L+1} \\ \vdots \quad \ddots & \vdots & \vdots \\ c_L \quad \cdots & c_{L+M-1} & c_{L+M} \\ 0^M \quad \cdots & 0 & 1 \end{vmatrix} = \begin{vmatrix} c_{L-M+1} \cdots & c_L \\ \vdots \quad \ddots & \vdots \\ c_L \quad \cdots & c_{L+M-1} \end{vmatrix}$$

Matrices (determinants resp.) of this type — that is with antidiagonals all constant — is called a Hankel matrices (Hankel determinants resp.). Since the Hankel determinants give a useful form for expressing some conditions of interest we will define $H(n, m)$ to be the determinant of the $(m + 1) \times (m + 1)$ Hankel matrix with $c_n, \dots, c_{n+2m}$ in the antidiagonals,

$$H(n, m) := \begin{vmatrix} c_n & \cdots & c_{n+m} \\ \vdots & \ddots & \vdots \\ c_{n+m} & \cdots & c_{n+2m} \end{vmatrix}.$$

We can relate the two previous equations by

$$Q^{[L/M]}(0) = H(L - M + 1, M - 1)$$

so that the sufficient condition for the existence of $R^{[L/M]}(z)$ is $H(L-M+1, M-1) \neq 0$.

We now narrow our focus to Padé approximants to a Hamburger transform. Let μ be a Borel measure on $\mathbb{R}$ with finite moments $c_k = \int p^k \, d\mu$.

With the shape reconstruction application in mind, we will now assume the measure μ is absolutely continuous with respect to the Lebesgue measure and denote it $f(x)dx$ where $f(x)$ is some non-negative Lebesgue measurable function of $\mathbb{R}$. Note that by this assumption, μ cannot be supported on a finite set. The additional restriction that μ has infinite support turns out to be important when discussing the determinacy of μ.

Recall that the Hamburger transform of μ,

$$g(z) := \int_{\mathbb{R}} \frac{d\mu}{1+pz}$$

has the asymptotic expansion, in the sense of non-tangential limits,

$$g(z) \simeq \sum_{k=0}^{\infty} c_k(-z^k).$$

We call this formal power series a Hamburger series.

Lemma 1.7.3. The Hamburger moments c_k satisfy the determinental condition $H(2n, m) \neq 0$ for $n, m = 0, 1, \ldots$. That is,

$$\begin{vmatrix} c_{2n} & \cdots & c_{2n+m} \\ \vdots & \ddots & \vdots \\ c_{2n+m} & \cdots & c_{2n+2m} \end{vmatrix} \neq 0$$

Proof. Consider the bilinear quadratic form given by

$$G(\mathbf{x}, \mathbf{y}) := \mathbf{x}^\top \begin{pmatrix} c_{2n} & \cdots & c_{2n+m} \\ \vdots & \ddots & \vdots \\ c_{2n+m} & \cdots & c_{2n+2m} \end{pmatrix} \mathbf{y}$$

If $\mathbf{x} = (x_0, x_1, \ldots, x_m)^\top$ then

$$\begin{aligned} G(\mathbf{x}, \mathbf{x}) &= \sum_{i,j=0}^{m} x_i x_j c_{2n+i+j} \\ &= \sum_{i,j=0}^{m} x_i x_j \int p^{2n+i+j}\, d\mu \\ &= \int \sum_{i,j=0}^{m} x_i x_j p^{2n+i+j}\, d\mu \\ &= \int p^{2n} \left(\sum_{k=0}^{m} x_k p^k \right)^2 d\mu \geq 0. \end{aligned}$$

Thus $G(\mathbf{x}, \mathbf{y})$ is positive semi-definite. Furthermore, equality holds

$$\int p^{2n} \left(\sum_{k=0}^{m} x_k p^k \right)^2 d\mu = 0$$

if and only if μ is supported entirely on the zeros of the polynomial $p^n \sum_{k=0}^{m} x_k p^k$. However by assumption μ has infinite support. So in fact G is strictly positive definite. We conclude, by Sylvester's criterion for example, that associated Hankel matrix has positive determinant. That is, $H(2n, m) > 0$. $\square$

The previous lemma guarantees the existence of certain Padé approximants; specifically those with $L - M + 1$ even.

Theorem 1.7.4. If J is odd then the Padé approximant $R^{[L/L+J]}(z)$ to the Hamburger series exists in Baker's sense.

Let $R_M(z)$ be the "offdiagonal" approximant $R^{[M/M+1]}(z)$ to the Hamburger series of μ,

$$R_M(z) = \frac{P_M(z)}{Q_M(z)} \approx \sum_{k=0}^{\infty} c_k (-z)^k \simeq \int_{-\infty}^{\infty} \frac{f(p)}{1+zp}\, dp$$

where $c_k = \int p^k f(p) dp$. In this section we will prove

Theorem 1.7.5. The off-diagonal Padé approximants $R_M(z)$ to the a determinate Hamburger series $\sum_{k=0}^{\infty} c_k (-z)^k$ exist and converge to $g(z)$ uniformly on compact subsets of the upper half plane $\{\mathrm{Im}\ z > 0\}$.

An outline of the proof is as follows: We first justify that the approximants exist in Baker's sense. Then it can be shown that limit of any convergent subsequence of R_M must have a representation as the Hamburger transform of some Borel measure μ which is a solution to our moment problem, and that furthermore such convergent subsequences exist. Thus in order for the sequence R_M to converge to $g(z)$ it will be necessary and sufficient that the moment problem be determinate.

Proposition 1.7.6. If the sequence c_k gives a determinate moment problem then the off-diagonal Padé approximants $R_M(z)$ converge to $g(z)$ locally uniformly.

Proof. The Padé approximant $R_M(z)$ has a convenient representation as an inner product in terms of a finite Jacobi matrix,

$$R_M(z) = \langle \delta_0, (1 + zA_M)^{-1} \delta_0 \rangle.$$

Now since A_M is a real matrix and $\frac{1}{|w|} \geq \frac{1}{|\mathrm{Im}(w)|}$ for any $w \in \mathbb{C}$, we see that

$$|zR_M(z)| = |\langle \delta_0, (z^{-1} + A_M)^{-1} \delta_0 \rangle| \leq \frac{1}{|\mathrm{Im}(1/z)|}$$

and thus

$$|R_M(z)| \leq \frac{1}{|z||\mathrm{Im}(1/z)|} = \frac{|z|}{|\mathrm{Im}(z)|}.$$

Since this bound is independent of R_M, Montel's theorem implies that the off-diagonal Pade approximants form a normal family. It can be shown that the limit of any convergent subsequence has a representation $\int (1 + xz) d\sigma(x)$ where σ is a solution to the Hamburger moment problem. Since the moment problem is determinate, the sequence R_M must converge to $g(z)$ uniformly on compact sets. $\square$

It remains to discuss the determinacy of the Hamburger moment problem. Here we need to add an additional constraint on the measure $d\mu = f(p)dp$, which is that $f(p)$ is L^2 integrable with respect to the Gaussian weight.

Proposition 1.7.7. A function $f(p) \in L^2(\mathbb{R}, e^{-p^2}dp)$ such that the moments

$$c_k = \int_{-\infty}^{\infty} p^k f(p) e^{-p^2}\ dp,\ \ k = 0, 1 \dots \tag{1.7.4}$$ eq:3

are finite, is uniquely determined by those moments.

Proof. It is sufficient to show that if $c_k = 0$ for all $k \geq 0$, then $f \equiv 0$ a.e. The Hermite moments are just linear combinations of 0,

$$\int_{-\infty}^{\infty} H_k(p) f(p) e^{-p^2}\ dp = 0,\ \ k \geq 0$$

where $H_k(p)$ is the Hermite polynomial of order k. Since the Hermite polynomials are complete in the space $L^2(\mathbb{R}, e^{-p^2}dp)$ then $f \equiv 0$. $\square$

1.8 Symmetric tensors and polynomials

Let V be a finite dimensional real vector space with basis $\{e_1, e_2, \dots, e_n\}$. The tensor power $V^{\otimes k}$ is a real vector space, with basis

$$\{e_{i_1} \otimes e_{i_2} \otimes \cdots \otimes e_{i_k}\}_{(i_1,\dots i_k) \in \{1,\dots n\}^k}$$

By convention we order this basis lexicographically. For example in the case $n = 3$, $k = 2$,

$$\begin{aligned}\{&e_1 \otimes e_1,\ e_1 \otimes e_2,\ e_1 \otimes e_3,\\ &e_2 \otimes e_1,\ e_2 \otimes e_2,\ e_2 \otimes e_3,\\ &e_3 \otimes e_1,\ e_3 \otimes e_2,\ e_3 \otimes e_3, \}\end{aligned}$$

For convenience we will use the vector notation $\mathbf{e} = (e_1, \dots, e_n)$ for the basis of V and

$$\mathbf{e}_{\otimes i} = e_{i_1} \otimes e_{i_2} \otimes \cdots \otimes e_{i_k}$$

for the elements of the induced basis of $V^{\otimes k}$, where $i = (i_1, i_2, \dots i_k) \in \{1, 2, \dots, n\}^k$. There is a natural embedding $V^k \hookrightarrow V^{\otimes k}$ given by

$$(v_1, v_2, \dots, v_k) \mapsto v_1 \otimes v_2 \otimes \cdots \otimes v_k$$

and the tensor power satisfies the following universal property: Any "multilinear" map out of V^k can be factored as a linear map out of $V^{\otimes k}$ composed with the inclusion above.

Commutative diagram NEEDED

Given a linear transformation $T : V \to V$, the tensor power $T^{\otimes k}$ is the unique linear map $V^{\otimes k} \to V^{\otimes k}$ such that

$$T^{\otimes k}(\mathbf{e}_{\otimes i}) = (Te_{i_1}) \otimes \cdots \otimes (Te_{i_k})$$

In the basis $\{\mathbf{e}_{\otimes i}\}$ ordered lexicographically the matrix representation for $T^{\otimes k}$ is given explicitly by the k-fold Kronecker product of the matrix for T in the basis $\mathbf{e}$. Tensor powers of linear maps commute with transposes

$$(T^{\otimes k})^\top = (T^\top)^{\otimes k}$$

If V has an inner product $\langle \cdot, \cdot \rangle_V$ then an inner product on $V^{\otimes k}$ can be defined by

$$\langle \mathbf{e}_{\otimes i}, \mathbf{e}_{\otimes i'} \rangle_{V^{\otimes k}} = \left\langle e_{i_1}, e_{i'_1} \right\rangle_V \cdots \left\langle e_{i_k}, e_{i'_k} \right\rangle_V .$$

We will generally omit the subscripts and write $\langle \cdot, \cdot \rangle$ for either inner product, determined from context. For any linear transformation $T : V \to V$ we have

$$\left\langle T^{\otimes k}(\mathbf{e}_{\otimes i}), \mathbf{e}_{\otimes i'} \right\rangle = \left\langle \mathbf{e}_{\otimes i}, (T^\top)^{\otimes k}(\mathbf{e}_{\otimes i'}) \right\rangle .$$

In particular for an orthogonal transformation $R \in O(n)$ on $\mathbb{R}^n$ we have

$$\left\langle R^{\otimes k}(\mathbf{e}_{\otimes i}), \mathbf{e}_{\otimes i'} \right\rangle = \left\langle \mathbf{e}_{\otimes i}, (R^{-1})^{\otimes k}(\mathbf{e}_{\otimes i'}) \right\rangle$$

and for a projection P,

$$\left\langle P^{\otimes k}\mathbf{e}_{\otimes i}, \mathbf{e}_{\otimes i'} \right\rangle = \left\langle \mathbf{e}_{\otimes i}, P^{\otimes k}\mathbf{e}_{\otimes i'} \right\rangle .$$

Note that

$$\left\langle T^{\otimes k}(\mathbf{e}_{\otimes i}), \mathbf{e}_{\otimes i'} \right\rangle = \left\langle Te_{i_1}, e_{i'_1} \right\rangle \cdots \left\langle Te_{i_k}, e_{i'_k} \right\rangle = T_{i_1, i'_1} \cdots T_{i_k, i'_k}$$

is a product of entries in the matrix representation for T with basis $\mathbf{e}$.

The inner product allows for a convenient expression of polynomials. Henceforth we take $V = \mathbb{R}^n$. For $x \in \mathbb{R}^n$ a multi index $\alpha \in \mathbb{N}_0^n$ of degree $|\alpha| = k$ we have

$$x^\alpha = x_1^{\alpha_1} x_2^{\alpha_2} \cdots x_n^{\alpha_n} = \prod_{i=1}^{n} \langle x, e_i \rangle^{\alpha_n} = \left\langle x^{\otimes k}, \mathbf{e}^{\otimes \alpha} \right\rangle$$

where we write

$$x^{\otimes k} = \underbrace{x \otimes x \otimes \cdots \otimes x}_{k \text{ times}}$$

and

$$\begin{aligned}\mathbf{e}^{\otimes\alpha} &= e_1^{\otimes\alpha_1} \otimes \cdots \otimes e_n^{\otimes\alpha_n} \\ &= \underbrace{e_1 \otimes \cdots \otimes e_1}_{\alpha_1 \text{ times}} \otimes \cdots \otimes \underbrace{e_n \otimes \cdots \otimes e_n}_{\alpha_n \text{ times}}\end{aligned}$$

For example, in $\mathbb{R}^3$ the monomial $x^{(1,1,2)} = x_1 x_2 x_3^2$ can be written

$$x_1 x_2 x_3^2 = \langle x \otimes x \otimes x \otimes x, e_1 \otimes e_2 \otimes e_3 \otimes e_3 \rangle$$

Now this monomial representation is not unique. Indeed, since the tensor inner product is given as a product of euclidean products, we may as well imagine $\mathbf{e}^{\otimes\alpha}$ to be commutative. More precisely, for any permutation $\sigma \in S_k$ we have equivalently,

$$\left\langle x^{\otimes k}, \mathbf{e}_{\otimes i} \right\rangle = \left\langle x^{\otimes k}, e_{\sigma^{-1} i_1} \otimes \cdots \otimes e_{\sigma^{-1} i_k} \right\rangle$$

for any $i \in \{1, \ldots, n\}^k$. The basis $\{\mathbf{e}_{\otimes i}\}_{i \in \{1,\ldots,n\}^k}$ can partitioned by equivalence under permutation to $\mathbf{e}^{\otimes\alpha}$ for each degree k multiindex α. Let

$$I(\alpha) = \{i \in \{1, \ldots, n\}^k : \left\langle x^{\otimes k}, \mathbf{e}_{\otimes i} \right\rangle = \left\langle x^{\otimes k}, \mathbf{e}^{\otimes\alpha} \right\rangle\}$$

One can check that $\{I(\alpha)\}_{|\alpha|=k}$ is a partition of the k-tuples $\{1, \ldots, n\}^k$. For example with $n = 2, k = 2$, we have

$$\begin{aligned} x_1^2 &= \langle x \otimes x, e_1 \otimes e_1 \rangle \\ x_1 x_2 &= \langle x \otimes x, e_1 \otimes e_2 \rangle = \langle x \otimes x, e_2 \otimes e_1 \rangle \\ x_2^2 &= \langle x \otimes x, e_2 \otimes e_2 \rangle \end{aligned}$$

Thus it seems more appropriate to write monomials in terms of "commutative" symmetric tensors. We define the symmetrization of a basis tensor $\mathbf{e}_{\otimes i}$ by averaging over permutations

$$P_s \mathbf{e}_{\otimes i} := \frac{1}{k!} \sum_{\sigma \in S_k} e_{\sigma i_1} \otimes \cdots \otimes e_{\sigma i_k}.$$

which can be seen as an orthogonal projection onto the subspace of symmetric tensors (tensors invariant under permutation).

The subspace of k-tensors invariant under permutation of indices is denoted $V^{\odot k} \subseteq V^{\otimes k}$. Elements may be defined by the projection,

$$v_1 \odot v_2 \odot \cdots \odot v_k = \frac{1}{k!} \sum_{\sigma \in S_k} v_{\sigma 1} \otimes v_{\sigma 2} \otimes \cdots \otimes v_{\sigma k}.$$

The non-decreasing tuples $i = (i_1, \ldots i_k) \in \{1, \ldots, n\}^k$ indexes a standard basis,

$$\mathbf{e}_{\odot i} := e_{i_1} \odot \cdots \odot e_{i_k}, \qquad 1 \leq i_1 \leq \cdots \leq i_k \leq n.$$

Even better, there is a one to one correspondence between non-decreasing k-tuples and multi-indeces of degree k,

$$\tilde{\alpha} = (\underbrace{1, \dots, 1}_{\alpha_1 \text{ times}}, \underbrace{2, \dots, 2}_{\alpha_2 \text{ times}}, \dots, \underbrace{n, \dots, n}_{\alpha_n \text{ times}})$$

so we may write the basis

$$\mathbf{e}^{\odot\alpha} := e_1^{\odot\alpha_1} \odot \cdots \odot e_n^{\odot\alpha_n}, \qquad \alpha \in \mathbb{N}_0^n, |\alpha| = k$$

The k-tensor inner product restricted to $V^{\odot k}$ says

$$\begin{aligned} \langle v_1 \odot \cdots \odot v_k, w_1 \odot \cdots \odot w_k \rangle &= \sum_{\sigma_1, \sigma_2 \in S_k} \prod \langle v_{\sigma_1 j}, w_{\sigma_2 j} \rangle \\ &= \sum_{\sigma \in S_k} \prod \langle v_{\sigma j}, w_j \rangle \\ &= \sum_{\sigma \in S_k} \prod \langle v_, w_{\sigma j} \rangle \,. \end{aligned}$$

Now the monomial expression in terms of symmetric tensors takes the simple form

$$x^\alpha = \left\langle x^{\odot k}, \mathbf{e}^{\odot\alpha} \right\rangle .$$

Taken further, the correspondence between homogeneous polynomials of degree k and symmetric k-tensors can be shown to be a vector space isomorphism, and extended to an isomorphism between the entire polynomial space and the graded space formed by symmetric tensors of all orders.

In the end one might wonder, if symmetric tensors are just polynomials in disguise, why bother with them?

Chapter 2

Moments of the RT and GRT

In this chapter we prove a basic result on the determinacy of certain multivariate moment problems by way of the Gaussian Radon transform.

Proposition 2.0.1. Let $c_\alpha(\omega) = \int_{-\infty}^{\infty} R_f(\omega, p)p^k dp$ be the projection moments of f at a fixed ω, and c_α the multivariate moments of f. Then

$$c_k(\omega) = \sum_{|\alpha|=k} \binom{k}{\alpha} \omega^\alpha c_\alpha$$

where $\binom{k}{\alpha} = \frac{k!}{\alpha_1!\alpha_2!\cdots\alpha_n!}$ are multinomial coefficients.

Proof. By the slice theorem (eq:ST 1.4.2) with $F(p) = p^k$,

$$\int_{-\infty}^{\infty} R_f(\omega, p)p^k \, dp = \int_{\mathbb{R}^n} f(x)\langle x, \omega\rangle^k \, dx.$$

Now $\langle x, \omega\rangle^k = (x_1\omega_1 + \cdots + x_n\omega_n)^k$ has the multinomial expansion

$$\langle x, \omega\rangle^k = \sum_{|\alpha|=k} \binom{k}{\alpha} x^\alpha \omega^\alpha.$$

Thus after a bit of rearranging we get

$$\begin{aligned} \int_{\mathbb{R}^n} f(x)\langle x, \omega\rangle^k \, dx &= \int_{\mathbb{R}^n} f(x) \sum_{|\alpha|=k} \binom{k}{\alpha} x^\alpha \omega^\alpha \, dx \\ &= \sum_{|\alpha|=k} \binom{k}{\alpha} \omega^\alpha \int_{\mathbb{R}^n} f(x) x^\alpha \, dx, \end{aligned}$$

where the integrands are precisely the kth degree multivariate moments of f. □

Similarly, moments of the GRT (Gaussian projection moments) can be expressed in terms of multivariate gaussian moments.

Proposition 2.0.2. Let $c_k^G(\omega) = \int_{-\infty}^{\infty} GR_f(\omega, p)p^k w(p)dp$ be the Gaussian moments of the GRT of f at a fixed ω. Let $c_\alpha^G = \int_{\mathbb{R}^n} f(x)w_n(x)x^\alpha dx$ be the Gaussian multivariate moments of f. Then

$$c_k^G(\omega) = \sum_{|\alpha|=k} \binom{k}{\alpha} \omega^\alpha c_\alpha^G.$$

Proof. The proof follows as it did for the RT. This time we apply the GRT slice theorem (1.4.3) with $F(p) = p^k$,

$$\int_{-\infty}^{\infty} GR_f(\omega, p) p^k w(p)\ dp = \int_{\mathbb{R}^n} f(x) \langle x, \omega \rangle^k w_n(x)\ dx$$

Again we use the multinomial expansion of $\langle x, \omega \rangle^n$ and rearrange:

$$\int_{\mathbb{R}^n} f(x) \langle x, \omega \rangle^k w_n(x)\ dx = \sum_{|\alpha|=k} \binom{k}{\alpha} \omega^\alpha \int_{\mathbb{R}^n} f(x) w_n(x) x^\alpha dx.$$

Thus

$$c^G(\omega) = \sum_{|\alpha|=k} \binom{k}{\alpha} \omega^\alpha c_\alpha^G.$$

□

Example 2.0.3. Let $e_1, e_2, \ldots, e_n \in S^{n-1}$ be the the standard basis for $\mathbb{R}^n$,

$$(1, 0, \ldots, 0),\ (0, 1, \ldots, 0),\ \ \ldots,\ (0, 0, \ldots, 1)$$

Then $\langle x, e_i \rangle = x_i$ is the natural projection of $\mathbb{R}^n$ onto the e_i axis. The standard projection moments can be calculated as follows

$$\begin{aligned} c_k(e_i) &= \sum_{|\alpha|=k} \binom{k}{\alpha} e_i^\alpha c_\alpha \\ &= c_{ke_i} \end{aligned}$$

since $e_i^\alpha = 0$ unless $\alpha = ke_i$.

The following theorem, due to Petersen [**CITATION NEEDED**], gives a way us to reduce the question of determinacy for multivariate moment problems to the classical case.

Proposition 2.0.4 (Petersen's theorem)**.** Let μ be a Borel measure with finite moments on $\mathbb{R}^n$, and $e_1, \ldots, e_n$ the standard basis for $\mathbb{R}^n$. If each $R_\mu^{e_1}, \ldots, R_\mu^{e_n}$ is determinate, then μ is determinate.

Proof. An outline of the proof is as follows. Preliminarily, note that the solution set $[\mu]$ of Borel measures with equivalent moments to μ is convex [**CITATION NEEDED**], and a μ is an extreme point in $[\mu]$ if and only if polynomials are dense in $L^1(\mathbb{R}^n, \mu)$ [**CITATION NEEDED**]. Thus it suffices to show that polynomials are dense in $L^1(\mathbb{R}^n, \mu')$ for any $\mu' \in [\mu]$.

The family of products of continuous functions of compact support $f(x) = \prod_{i=1}^n f_i(x_i)$ where each $f_i \in C_c(\mathbb{R})$, is dense in $L_1(\mu)$ [**CITATION NEEDED**]. Furthermore, since $R_\mu^{e_i}$, $i = 1, \ldots, n$ are determinate, polynomials are dense in each $L^2(\mathbb{R}, R_\mu^{e_i})$ [**CITATION NEEDED**]. Petersen constructs a series of polynomials $P_i : \mathbb{R} \to \mathbb{R}$ such that the polynomial product $P(x) = \prod_{i=1}^n P_i(x_i)$ arbitrarily close to f in $L^1(\mu)$. Thus polynomials are dense in $L^1(\mathbb{R}^n, \mu)$ and the moment problem is determinate. [**PROOF NEEDED**]

Note Schmüdgen proves this via Borel characteristic functions. Not clear what the benefit is. □

Remark 2.0.5. Conjecture? We expect this result to hold true for any orthonormal basis, and likely any basis. Will have to check this. Arguable benefit for us is that we in theory can use any n linearly independent projections to reconstruct μ.

Corollary 2.0.6. If μ is compactly supported Borel measure on $\mathbb{R}^n$, then the multivariate moment problem is determinate.

Proof. It suffices to note that the projections $R_\mu^{e_i}$ are compactly supported Borel measures on $\mathbb{R}$, and thus determinate by [**CITATION NEEDED**]. □

We show in section (2.2) that any function f in the weighted space $L^2(\mathbb{R}, \gamma)$ is determinate (maybe this should be moved to section 1.2). Petersen's theorem allows us to generalize this result to $\mathbb{R}^n$.

Corollary 2.0.7. If $f \in L^2(\mathbb{R}^n, \gamma^n)$, then f is determinate.

Proof. [**PROOF NEEDED**] □

Chapter 3

Shape reconstruction from moments

In this chapter we describe a method for shape reconstruction from moments and propose an extension of the method to unbounded regions.

3.1 Convergence results for the Gaussian Radon transform

Take a non-negative Lebesgue measurable function $f(x)$ on $\mathbb{R}^n$ with finite multivariate Gaussian moments

$$c_\alpha^G = \int_{\mathbb{R}^n} f(x) e^{-\|x\|^2/2} x^\alpha dx$$

We have seen that the multivariate Stieltjes transform of $f(x)w_n(x)$,

$$g(z) = \int_{\mathbb{R}^n} \frac{f(x)e^{-\|x\|^2/2}}{1 + \langle x, \omega \rangle z} dx,$$

is equal to the Hamburger transform of $GR_f(\omega, p)$, whose moments in turn are easily computed from the moments of f, for fixed ω. Thus for z in the one dimensional subspace spanned by ω, $g(z)$ can be approximated well by Padé approximants, under the integrability condition $GR_f(\omega, p) \in L^2(\mathbb{R}, e^{-p^2}\ dp)$. On the other hand as an integral on $\mathbb{R}^n$, we can also approximate $g(z)$ by cubature formula,

$$\int_{\mathbb{R}^n} \frac{f(x)e^{-\|x\|^2/2}}{1 + \langle x, \omega \rangle z} dx \approx \sum_{\ell=L} \frac{e^{-\|x_\ell\|^2/2}}{1 + \langle x_\ell, \omega \rangle z} w_\ell f(x_\ell) =: \sum_{\ell=L} C_\ell(z) f(x_\ell).$$

By equating these parallel approximations,

$$\sum_{\ell=L} C_\ell(z) f(x_\ell) \approx R_M(z),$$

with a sufficient quantity of sample points z_j we have arrived at a linear system

$$\sum_{\ell=L} C_\ell(z_j) f(x_\ell) = R_M(z_j), \qquad j = 1, 2, \ldots, J$$

from which we should be able to recover the value of f on our cubature nodes x_ℓ.

3.2 Implementation

We have spent some time implementing our proposed extension to the shape reconstruction method, mostly in the Wolfram Mathematica language. But of course in this process questions arose, both theoretical and pragmatic.

While the authors describe forming multivariate "homogeneous" Pade approximants — perhaps first proposed by Cuyt herself in [?Cuyt84] — we opt to bypass this step using univariate approximants instead. Indeed, the only benefit to Cuyt's approximants explicitly cited by the authors, as compared to other multivariate generalizations of the Pade approximants, is that they are "the only satisfying the following powerful slice theorem" [?Cuyt05]. This slice theorem simply states that the homogeneous approximants when restricted to one dimensional subspaces, are equivalent to univariate Pade approximants. Since we only need a finite number of sample points it seems reasonable to compute a series of univariate approximants at various projection angles.

The decision to drop the homogenous approximants has the following potential consequences: On the one hand, the univariate approximants are simpler to compute, with methods built in to many CASs. On the other, while the authors select sample points from a cubic lattice, it would be inefficient (though potentially less so with careful selection of projection angles via, say, Farey sequences) to do so with one dimensional subspaces. Instead we take a collection of equidistant projection angles forming a radial grid of sample points. This seems acceptable, as it is unclear if there is any benefit to the sample points having a particular geometric structure. We should also agknowledge the potential that forming a single homogeneous appoximant is more computationally efficient than many univariate approximants. This requires further investigation, but we note that (anecdotally) in computational tests the formation of approximants is not very intensive compared to the subsequent solving of the high dimensional linear system. Some thought was given to the possibility that sampling from a cubic lattice provides some structure which makes the system easier to solve, but so far we have no hard evidence of this.

A couple of notes on computational details: First we should be explicit about the line of computation. At the outset, the given information is a certain multivariate moment sequence (up to a finite order with some accuracy assumption) which we assume belongs to a region within some fixed bounding box. In evaluating the computational viability of the method we take note of which steps depend on the moments, and which can be precomputed. In order to form univariate approximants we first compute projection moments. While these of course depend on the moment sequence, the multinomial formula used for this computation only depends on the projection angle. Thus with a preditermined set of projection angles this is partially precomputed. Since the cubature formula itself only depends on the choice of quadrature nodes and sample points, it can be precomputed and applied to any moment sequence. The choice of quadrature formula is somewhat arbitrary. Lacking expertise in the area of numerical integration, we follow the authors, opting for the four-point Gauss-Legendre product formula. This formula takes nodes from a nearly cubic lattice, which is implicitly quantized when forming the pixel image. We have not evaluated the error introduced by this quantization but expect it to be minimal and to approach zero with higher pixel resolution.

We don't expect the linear system to have an exact solution, let alone one that whose computation is tractable for high resolutions. We may try Mathematica's LeastSquares method, but Cuyt et al. specifically mention using a truncated singular value decomposition.

For my own benefit, let's review singular value decompositions. An $i \times j$ matrix A of rank r can be written as

$$A = U\Sigma V^\top,$$

where Σ is the diagonal matrix of singular values $\sigma_1, \sigma_2, \ldots, \sigma_r$, and U and V are unitary matrices. The truncated singular value decomposition gives us a reduced rank approximation to A by replacing all but the largest (first) k singular values in Σ with zeros,

$$\tilde{\Sigma} = \mathrm{diag}(\sigma_1, \sigma_2, \ldots, \sigma_k, 0, \ldots, 0)$$

The truncated SVD matrix $\tilde{A} = U\tilde{\Sigma}V^\top$ is optimal in the sense that it is the closest rank-k matrix to A in the Frobenius norm.

The SVD and truncated SVD are used to solve the linear least squares problem as follows: From $U\Sigma V^\top f = p$ we get $f = V\Sigma^{-1}U^\top p$, and similarly $f = \tilde{V}\tilde{\Sigma}^{-1}U^\top p$. Note that the psuedoinverse $\tilde{\Sigma}^{-1}$ is a diagonal matrix made up of the reciprocal singular values.

Mathematica offers a few built in functions for this. LeastSquares packages a number of methods to solve a linear least squares based on the sparsity of the System (our system is dense). SingularValueDecomposition can be restricted to a certain number of singular values. PsuedoInverse can specify a "tolerance" zeroing out singular values less than some proportion of the maximum singular value. While LeastSquares is the simplest option, it does not seem to support truncated solutions. Between SingularValueDecomposition and PsuedoInverse, both can handle truncated problems, but the latter seems — at least superficially — better suited for out application.

On Gaussian shape reconstruction. There remain some minor practical concerns about, for example, the choice of quadrature formula and sample points. An obvious choice would be to continue with the Gauss-Legendre product formula, now applied on a Gaussian-weighted f. Some consideration should be taken as to the viability of other Gaussian quadrature formulas defined particulary for the Gaussian measure.

Now, a fundamental limitation with the proposed method which we have neglected to mention is the practical one: As a computational method we are necessarily restriced to the finite — finite moments, finite order approximations, and most problamatic, bounded domains. In what application would one even be interested in unbounded shape reconstruction, while also having access of Gaussian moments? It would be reasonable to question the use of a computational method in an unbounded context. As with any finite approximation, the unboundedness should be thought of as an idealized limit; perhaps an unbounded approximation is a limit of bounded approximations. But then, the original method applies just as well if all we wanted was a sequence of reconstructions expanding in range. We can only speculate at this point that perhaps this method, being taylored specifically for unbounded reconstruction, could hold some practical advantages over its predecessor; or be satisfied with impracticality.

3.3 An example

First compare the methods on bounded regions: Unit ball in $\mathbb{R}^3$? Annulus in $\mathbb{R}^2$? Some asymmetric or off-center examples?

More importantly, present examples for unbounded regions: Strips and cones come to mind. One-sided strips maybe.

How do we "image" an unbounded region in the first place? Our method places pixels upon the nodes of a cubature formula which is usually a bounded cubic lattice. What are appropriate nodes for, and how can we best "image" a region which does not fit within a bounded rectangle?

Chapter 4

The Gaussian Radon transform of polynomials

Unlike the Radon transform, the Gaussian Radon transform is well defined on polynomials. In this chapter we investigate the GRT from this perspective, as a linear operator between polynomial spaces. In particular we prove a simple formula for the GRT of multivariate Hermite polynomials. We extend this result to some generalizations of the GRT.

4.1 The Radon transform of multivariate Hermite polynomials over general affine subspaces

We begin this section by proving a formula for the GRT of the multivariate Hermite polynomials defined in the previous section. Recall that for $\alpha \in \mathbb{N}_0^n$ the polynomials $H_\alpha(x)$ can be defined by the generating function

$$e^{\langle x,y\rangle - \frac{\|y\|^2}{2}} = \sum_{\alpha \in \mathbb{N}_0^n} \frac{H_\alpha(x)}{\alpha!} y^\alpha$$

where $\alpha! = \alpha_1! \cdots \alpha_n!$. These are related to the classical Hermite polynomials by

$$H_\alpha(x) = \prod_{i=1}^{n} H_{\alpha_i}(x_i)$$

where, for $k \in \mathbb{N}_0^\infty$, the classical polynomials $H_k(p)$ can be defined by

$$e^{pt - \frac{t^2}{2}} = \sum_{k=0}^{\infty} \frac{H_k(p)}{k!} t^k.$$

prop:GRTHermite **Proposition 4.1.1.** Let $\omega \in S^{n-1}$ and $p \in \mathbb{R}$ be fixed. If $|\alpha| = k$, then the GRT of the multivariate Hermite polynomial H_α is

$$GR_{H_\alpha}(\omega, p) = H_k(p)\omega^\alpha \tag{4.1.1}$$ eq:GRH

Proof. Recall that the generating function $\phi(x) = e^{\langle x,y\rangle - \frac{\|y\|^2}{2}}$ converges absolutely for all $x, y \in \mathbb{R}^n$. Consider the GRT of ϕ,

$$GR_\phi(\omega, p) = \int\limits_{\langle x,\omega\rangle = p} \phi(x) w_{n-1}(x - p\omega)\ dx.$$

Immediately we can expand $\phi(x)$, interchanging integral and series, to see

$$\begin{aligned} GR_\phi(\omega, p) &= \sum_{\alpha\in\mathbb{N}_0^n} \frac{1}{\alpha!} y^\alpha \int\limits_{\langle x,\omega\rangle = p} H_\alpha(x) w_{n-1}(x - p\omega)\ dx \\ &= \sum_{\alpha\in\mathbb{N}_0^n} \frac{GR_{H_\alpha}(\omega, p)}{\alpha!} y^\alpha. \end{aligned} \tag{4.1.2}$$

On the other hand, we will be able to show that

$$GR_\phi(\omega, p) = \sum_{k=0}^{\infty} \sum_{|\alpha|=k} \frac{H_k(p)\omega^\alpha}{\alpha!} y^\alpha. \tag{4.1.3}$$

Thus, being careful to note that the series (4.1.2) and (4.1.3) converge absolutely, we can compare coefficients for the desired formula. In order to derive the second expansion we begin by translating our integral onto the linear subspace $\langle x, \omega\rangle = 0$, via the change of variables $x \mapsto x + p\omega$:

$$\begin{aligned} GR_\phi(\omega, p) &= \int\limits_{\langle x,\omega\rangle = 0} \phi(x + p\omega) w_{n-1}(x)\ dx \\ &= \int\limits_{\langle x,\omega\rangle = 0} e^{\langle x+p\omega, y\rangle - \frac{\|y\|^2}{2}} (2\pi)^{-\frac{n-1}{2}} e^{-\frac{\|x\|^2}{2}}\ dx \\ &= e^{p\langle \omega, y\rangle - \frac{\|y\|^2}{2}} \int\limits_{\langle x,\omega\rangle = 0} e^{\langle x, y\rangle - \frac{\|x\|^2}{2}} (2\pi)^{-\frac{n-1}{2}}\ dx \end{aligned} \tag{4.1.4}$$

Now in order to compute the right side integral we consider the orthogonal decomposition $y = y_\omega + y_{\omega^\perp}$, where

$$y_\omega = \langle y, \omega\rangle\omega \qquad y_{\omega^\perp} = y - y_\omega.$$

We make two observations about this decomposition of y: First,

$$\|y\|^2 = \|y_\omega\|^2 + \|y_{\omega^\perp}\|^2 = \langle y, \omega\rangle^2 + \|y_{\omega^\perp}\|^2 \tag{4.1.5}$$

and second, for x in the linear subspace $\langle x, \omega\rangle = 0$ we have

$$\langle x, y\rangle = \langle x, \omega\rangle\langle y, \omega\rangle + \langle x, y_{\omega^\perp}\rangle = \langle x, y_{\omega^\perp}\rangle.$$

This second observation allows us to solve the integral at the end of (4.1.4). By completing the square $\|x - y_{\omega^\perp}\| = \|y_{\omega^\perp}\|^2 - 2\langle x, y_{\omega^\perp}\rangle + \|x\|^2$, we have

$$\int\limits_{\langle x,\omega\rangle=0} e^{\langle x,y\rangle - \frac{\|x\|^2}{2}} (2\pi)^{-\frac{n-1}{2}}\ dx = e^{-\frac{\|y_{\omega^\perp}\|^2}{2}} \int\limits_{\langle x,\omega\rangle=0} e^{-\frac{\|y_{\omega^\perp}\|^2}{2} + \langle x, y_{\omega^\perp}\rangle - \frac{\|x\|^2}{2}} (2\pi)^{-\frac{n-1}{2}}\ dx$$
$$= e^{-\frac{\|y_{\omega^\perp}\|^2}{2}} \int\limits_{\langle x,\omega\rangle=0} e^{-\frac{\|x-y_{\omega^\perp}\|}{2}} (2\pi)^{-\frac{n-1}{2}}\ dx$$

which is just $e^{-\frac{\|p_{\omega^\perp}\|^2}{2}}$. It was important here to invoke the orthogonal projection $y_{\omega^\perp}$ so that the integrand becomes the translation of a standard Gaussian function on the hyperplane $\langle x, \omega\rangle = 0$. Now returning to (eq:GRTPhiExpansion21 4.1.4), we have

$$GR_\phi(\omega, p) = e^{p\langle \omega, y\rangle - \frac{\|y\|^2}{2} - \frac{\|y_{\omega^\perp}\|^2}{2}}$$
$$= e^{p\langle \omega, y\rangle - \frac{\langle y,\omega\rangle^2}{2}}$$

where we used (eq:GRTPhiExpansion22 4.1.5). This looks like the generating function for the univariate Hermite polynomials $H_k(p)$ with $t = \langle \omega, y\rangle$, so we expand

$$GR_\phi(\omega, p) = \sum_{k=0}^{\infty} \frac{H_k(p)}{k!} \langle \omega, y\rangle^k.$$

Finally we apply the multinomial expansion of $\langle \omega, y\rangle^k$, recalling the multinomial coefficients $\binom{k}{\alpha} = \frac{k!}{\alpha!}$. Therefore

$$GR_\phi(\omega, p) = \sum_{k=0}^{\infty} \sum_{|\alpha|=k} \frac{H_k(p)}{k!} \binom{k}{\alpha} y^\alpha \omega^\alpha$$
$$= \sum_{k=0}^{\infty} \sum_{|\alpha|=k} \frac{H_k(p)\omega^\alpha}{\alpha!} y^\alpha$$

as needed. □

Remark 4.1.2. The proof above was written by Dr. Sengupta. Recall that the Hermite polynomials H_α form a complete orthogonal basis for $L^2(\mathbb{R}^n, w_n)$. In this sense the formula (eq:GRH 4.1.1) completely defines the Gaussian Radon transform on this space. Indeed, it can be shown that any function $f \in L^2(\mathbb{R}^n, w_n)$ has an expansion

$$f(x) = \sum_{\alpha \in \mathbb{N}_0^n} a_\alpha^{(f)} H_\alpha(x)$$

where

$$a_\alpha^{(f)} = \int_{\mathbb{R}^n} f(x) H_\alpha(x) w_n(x)\ dx.$$

If this expansion converges nicely on Λ then

$$\begin{aligned}
GR_f(\Lambda) &= \int_\Lambda f(x) w_n(x-p_0)\ dx \\
&= \int_\Lambda \sum_{\alpha\in\mathbb{N}_0^n} a_\alpha^{(f)} H_\alpha(x) w_n(x-p_0)\ dx \\
&= \sum_{\alpha\in\mathbb{N}_0^n} a_\alpha^{(f)} \int_\Lambda H_\alpha(x) w_n(x-p_0)\ dx \\
&= \sum_{\alpha\in\mathbb{N}_0^n} a_\alpha^{(f)} GR_{H_\alpha}(\Lambda)
\end{aligned}$$

If Λ is a hyperplane $\langle x, \omega\rangle = p$ with $\omega \in S^{n-1}$ then this gives

$$\begin{aligned}
GR_f(\Lambda) &= \sum_{\alpha\in\mathbb{N}_0^n} a_\alpha^{(f)} GR_{H_\alpha}(\Lambda) \\
&= \sum_{\alpha\in\mathbb{N}_0^n} a_\alpha^{(f)} H_{|\alpha|}(p)\omega^\alpha \\
&= \sum_{k=0}^{\infty} H_k(p) \sum_{|\alpha|=k} a_\alpha^{(f)} \omega^\alpha
\end{aligned}$$

For details on modes convergence see (Dunkl or Thangavelu) [**CITATION NEEDED**].

In section 1.4 we defined the RT and GRT as the integrals of a function $f : \mathbb{R}^n \to \mathbb{R}$ over $n-1$ dimensional hyperplanes, the standard definitions. But there are many examples of geometric integral transforms, closely related to the hyperplane RT, which can be thought of as "generalized Radon transforms". In fact, the "Funk transform", which relates a function on the sphere S^3 to its integrals over great circles, was first introduced by Paul Funk in 1911, half a decade before Radon introduced the hyperplane RT.İn the remainder of this chapter we discuss the notion of the RT and GRT on general d-dimensional affine subspaces of $\mathbb{R}^n$, where $d = 0, \ldots, n$. For context, note that when $d = 1$, this is the so called "X-ray transform" which gets it's name from the direct application to radiology.

Let Λ_0 be a d-dimensional linear subspace of $\mathbb{R}^n$, and $v \in \mathbb{R}^n$. Then the translation of Λ_0 by p

$$\Lambda = v + \Lambda_0$$

is called an **affine subspace** of $\mathbb{R}^n$. The representation above is clearly not unique since adding any member of Λ_0 to v does not change Λ. However, given an affine subspace Λ we have a canonical choice for v as the closest point in Λ to the origin. This point exists and is unique since Λ is closed and convex. Furthermore the point v defined in this way must be orthogonal to any vector in the linear subspace Λ_0. That is, $v \in \Lambda_0^\perp$, where $\Lambda_0^\perp$ is the linear subspace

$$\Lambda_0^\perp = \{x \in \mathbb{R}^n : \langle x, y \rangle = 0 \text{ for all } y \in \Lambda\}$$

also known as the **orthogonal complement** of Λ_0. Thus we rephrase our definition as follows:

Definition 4.1.3. An **affine subspace** Λ of dimension d in $\mathbb{R}^n$ is defined uniquely by

$$\Lambda = v + \Lambda_0$$

where Λ_0 is a linear subspace of dimension d and $v \in \Lambda_0^\perp$. Sometimes we call Λ a "d-plane".

Remark 4.1.4. Note that the hyperplanes in the original definition of the RT are $(n-1)$-planes, and the lines in the "X-ray transform" are 1-planes.

Recall that if Λ_0 is a fixed linear subspace of dimension d, the orthogonal complements $\Lambda_0^\perp$ is a linear subspace of dimension $n - d$. For any $x \in \mathbb{R}^n$ there is a unique orthogonal decomposition $x = x_{\Lambda_0} + x_{\Lambda_0^\perp}$ where $x_{\Lambda_0} \in \Lambda_0$ and $x_{\Lambda_0^\perp} \in \Lambda_0^\perp$. This idea can be summarized by the expression

$$\mathbb{R}^n = \Lambda_0 \oplus \Lambda_0^\perp.$$

Sometimes it is convenient to have an explicit coordinate system on an affine subspace. If $u^{(1)}, \ldots, u^{(d)}$ is a orthonormal basis for Λ_0 then there is an isometric embedding of $\mathbb{R}^d$ into $\mathbb{R}^n$ with image Λ given by $x(t) : \mathbb{R}^d \to \Lambda$ where

$$x(t) = v + t_1 u^{(1)} + \cdots + t_d u^{(d)}, \qquad t \in \mathbb{R}^d.$$

By this embedding can define the Euclidean measure dx on Λ as the push-forward of the Lebesgue measure on $\mathbb{R}^n$, and the standard Gaussian measure on Λ as the push-forward of γ^n. Note the $x(t)$ is defined in such a way that $x(0) = v$ which is essential when defining the Gaussian measure on Λ.

Other times we prefer not to invoke a basis-dependent isometry, and while $x(t)$ clearly depends on the choice of $u^1, \ldots, u^{(d)}$, we should note that the Euclidean and Gaussian measures on Λ are independent of basis [**PROOF NEEDED**].

Now we can define the natural analogue of the RT for affine subspaces, which is sometimes called the ***d*-plane transform**. To be consistent with the hyperplane RT we will write $v = p\omega$, where $\omega \in S^{n-1} \cap \Lambda_o^\perp$ and $p \in \mathbb{R}$. Of course this notation is unique only up to the identification $(\omega, p) = (-\omega, -p)$.

Definition 4.1.5. Let $f : \mathbb{R}^n \to \mathbb{R}$. For any affine subspace $\Lambda = p\omega + \Lambda_0$, we define the **Radon Transform of f on** Λ by

$$R_f(\Lambda) = \int_\Lambda f(x)\ dx$$

and the **Gaussian Radon Transform of f on Λ** by

$$GR_f(\Lambda) = \int_{\Lambda} f(x) w_d(x - p\omega)\ dx$$

where d is the dimension of Λ_0.

Remark 4.1.6. Explicitly, if $x(t) : \mathbb{R}^d \to \Lambda$ is an Euclidean isometry then

$$R_f(\Lambda) = \int_{\mathbb{R}^d} f(x(t))\ dt$$

If furthermore $x(0) = p\omega$ then

$$GR_f(\Lambda) = \int_{\mathbb{R}^d} f(x(t)) w(t)\ dt$$

However keep in mind that these definitions are independent of the particular isometry [**PROOF NEEDED**].

In order to generalize the formula (eq:GRH 4.1.1) we begin with the integral of H_α over linear subspaces.

Recall that the GRT of a function f over a linear subspace Λ_0 can be written

$$GR_f(\Lambda_0) = \int\limits_{\mathbb{R}^d \times \{0\}^{n-d}} f(Rx) w_d(x)\ dx \tag{4.1.6}$$

eq:GRRot

where $R \in SO(n)$ is the a rotation sending $\mathbb{R}^d \subseteq \mathbb{R}^n$ to Λ_0. Thus we would like to investigate the action of rotations on certain functions of interest, namely monomials and Hermite polynomials.

Lemma 4.1.7. Let $f(x) = x^\alpha$ for some multi-index $\alpha \in \mathbb{N}_0^n$ of degree $|\alpha| = k$. Then the rotation $f(Rx)$ has the expansion

$$f(Rx) = \sum_{\substack{\beta \in \mathbb{N}_0^n \\ |\beta| = k}} c(R)^\alpha_\beta x^\beta. \tag{4.1.7}$$

eq:monRot

The coefficients $c(R)^\alpha_\beta$ are given by

$$c(R)^\alpha_\beta = \sum_{i \in I(\beta)} \prod_{j=1}^{n} \left\langle Re_{i_j}, e_{\alpha^j} \right\rangle$$

where $\alpha^j = \ell$ if $0 < j - \alpha_1 - \cdots \alpha_{\ell-1} \leq \alpha_\ell$ and

$$I(\beta) = \{i \in \{1, \ldots, n\}^k : 1, \ldots, n \text{ have multiplicity } \beta_1, \ldots \beta_n \text{ in } i\}$$

Note that the inner products $\left\langle Re_{i_j}, e_{\alpha^j} \right\rangle$ are matrix entries for R.

Proof. First one should note that $(Rx)^\alpha$ is a homogeneous polynomial of degree

k; a polynomial because R is a linear transformation, and homogeneous because $R \in SO(n)$ so that for $\lambda \in \mathbb{R}$,

$$(R\lambda x)^\alpha = (\lambda R x)^\alpha = \lambda^k (Rx^\alpha).$$

Using the tensor power notation from appendix A, we can write

$$(Rx)^\alpha = \prod_{i=1}^n \langle Rx, e_i \rangle^{\alpha_i} = \left\langle R^{\otimes k} x^{\otimes k}, \mathbf{e}^{\otimes \alpha} \right\rangle = \left\langle x^{\otimes k}, \left(R^{-1}\right)^{\otimes k} \mathbf{e}^{\otimes \alpha} \right\rangle$$

Now we expand over the basis $\{\mathbf{e}_{\otimes i}\}_{i \in \{1,\ldots,n\}^k}$, noting that each tuple $i \in \{1,\ldots,n\}^k$ is an anagram to exactly one multiindex of degree k. To be precise $\{I(\beta)\}_{|\beta|=k}$ is a partition of $\{1,\ldots,n\}^k$. Thus

$$\begin{aligned}
\left\langle x^{\otimes k}, \left(R^{-1}\right)^{\otimes k} \mathbf{e}^{\otimes \alpha} \right\rangle &= \sum_{\substack{\beta \in \mathbb{N}_0^n \\ |\beta| = k}} \sum_{i \in I(\beta)} \left\langle x^{\otimes k}, \mathbf{e}_{\otimes i} \right\rangle \left\langle \mathbf{e}_{\otimes i}, \left(R^{-1}\right)^{\otimes k} \mathbf{e}^{\otimes \alpha} \right\rangle \\
&= \sum_{\substack{\beta \in \mathbb{N}_0^n \\ |\beta| = k}} x^\beta \sum_{i \in I(\beta)} \left\langle \mathbf{e}_{\otimes i}, \left(R^{-1}\right)^{\otimes k} \mathbf{e}^{\otimes \alpha} \right\rangle .
\end{aligned}$$

That we have the non-tensor expression

$$\left\langle \mathbf{e}_{\otimes i}, \left(R^{-1}\right)^{\otimes k} \mathbf{e}^{\otimes \alpha} \right\rangle = \prod_{j=1}^n \left\langle e_{i_j}, R^{-1} e_{\alpha^j} \right\rangle = \prod_{j=1}^n \left\langle R e_{i_j}, e_{\alpha^j} \right\rangle$$

follows from the definition of the inner product on $\left(\mathbb{R}^n\right)^{\otimes k}$ and further notation from appendix A. $\square$

Remark 4.1.8. We could write the coefficient $c(R)^\alpha_\beta$ in a slightly different form. If we define the non-decreasing tuple $\tilde{\alpha}$ associated with a multi-index α by

$$\tilde{\alpha} = (\underbrace{1,\ldots,1}_{\alpha_1 \text{ times}}, \underbrace{2,\ldots,2}_{\alpha_2 \text{ times}}, \ldots, \underbrace{n,\ldots,n}_{\alpha_n \text{ times}})$$

then

$$c(R)^\alpha_\beta = \sum_{\sigma \in S_k} \prod_{j=1}^n R_{\sigma \tilde{\beta}_j, \tilde{\alpha}_j}.$$

Example 4.1.9. For a simple example take the monomial $f(x) = x_1^2 x_2$. Here $n = 2$, $\alpha = (2,1)$ and $k = 3$. The partition of $\{1,2\}^3$ over multi-indices $|(\beta_1, \beta_2)| = 3$ is given by

$$\begin{aligned}
I(3,0) &= \{(1,1,1)\} \\
I(2,1) &= \{(1,1,2),(1,2,1),(2,1,1)\} \\
I(1,2) &= \{(1,2,2),(2,1,2),(2,2,1)\} \\
I(0,3) &= \{(2,2,2)\}
\end{aligned}$$

Now note that $(\alpha^1, \alpha^2, \alpha^3) = (1, 1, 2)$. If $R \in SO(2)$ is a 2×2 rotation matrix with entries $R_{i,j} = \langle Re_i, e_j \rangle$ then the formula (eq:monRot 4.1.7) says

$$\begin{aligned}(Rx)^\alpha = {} & x_1^3 R_{1,1}^2 R_{1,2} \\ & + x_1^2 x_2 \left(R_{1,1}^2 R_{2,2} + 2R_{1,1}R_{2,1}R_{1,2}\right) \\ & + x_1 x_2^2 \left(2R_{1,1}R_{2,1}R_{2,2} + R_{2,1}^2 R_{1,2}\right) \\ & + x_2^3 R_{2,1}^2 R_{2,2}\end{aligned}$$

While tensor power notation is concise, this sort of representation is needed for computation.

Proposition 4.1.10. The GRT of a monomial $f(x) = x^\alpha$ on a linear subspace Λ_0 is given by

$$GR_f(\Lambda_0) = \sum_{\substack{\beta \in \mathbb{N}_0^n \\ |\beta| = k}} c(R)_\beta^\alpha c_\beta^G$$

where $c_\beta^G = \int_{\mathbb{R}^d} x^\beta w_d(x)\ dx$ are the Gaussian moments on $\mathbb{R}^d$. **This should be indexed by only the first d entries of β. Feels weird**.

Proof. We take (eq:GRRot 4.1.6),

$$GR_f(\Lambda_0) = \int\limits_{\mathbb{R}^d \times \{0\}^{n-d}} (Rx)^\alpha w_d(x)\ dx,$$

and expand by (eq:monRot 4.1.7),

$$\int\limits_{\mathbb{R}^d \times \{0\}^{n-d}} (Rx)^\alpha w_d(x)\ dx = \sum_{\substack{\beta \in \mathbb{N}_0^n \\ |\beta| = k}} c(R)_\beta^\alpha \int_{\mathbb{R}^d \times \{0\}^{n-d}} x^\beta w_d(x)\ dx$$

as needed. $\square$

Proposition 4.1.11. The GRT of a multivariate Hermite polynomial H_α on an affine subspace Λ is given by

$$GR_{H_\alpha}(\Lambda) = \alpha! \sum_{\substack{\beta \in \mathbb{N}_0^n \\ |\beta| = |\alpha|}} c(P_{\Lambda_0^\perp})_\alpha^\beta \frac{H_\beta(p_0)}{\beta!}$$

Proof. Once again we take the GRT of the generating function $\phi(x) = e^{\langle x, y \rangle - \frac{\|y\|^2}{2}}$, initially expanding

$$GR_\phi(\Lambda) = \sum_{\alpha \in \mathbb{N}_0^n} \frac{GR_{H_\alpha}(\Lambda)}{\alpha!} y^\alpha$$

Simultaneously we have

$$
\begin{aligned}
GR_\phi(\Lambda) &= e^{\langle p_0, y\rangle - \frac{\|y\|^2}{2}} (2\pi)^{-\frac{d}{2}} \int_\Lambda e^{\langle x, y\rangle - \frac{\|x\|^2}{2}} \, dx \\
&= e^{\langle p_0, y\rangle - \frac{\|y\|^2}{2} + \frac{\left\|P_{\Lambda_0} y\right\|^2}{2}} \\
&= e^{\langle p_0, y\rangle - \frac{\left\|P_{\Lambda_0^\perp} y\right\|^2}{2}} \\
&= \sum_{\alpha \in \mathbb{N}_0^n} \frac{H_\alpha(p_0)}{\alpha!} (P_{\Lambda_0^\perp} y)^\alpha
\end{aligned}
$$

Now we should have noted that the formula (eq:monRot 4.1.7) applies not only to rotations but to any linear transformation, so we have

$$
\begin{aligned}
GR_\phi(\Lambda) &= \sum_{\alpha \in \mathbb{N}_0^n} \frac{H_\alpha(p_0)}{\alpha!} \sum_{\substack{\beta \in \mathbb{N}_0^n \\ |\beta| = |\alpha|}} c(P_{\Lambda_0^\perp})_\beta^\alpha y^\beta \\
&= \sum_{\beta \in \mathbb{N}_0^n} y^\beta \sum_{\substack{\alpha \in \mathbb{N}_0^n \\ |\alpha| = |\beta|}} c(P_{\Lambda_0^\perp})_\beta^\alpha \frac{H_\alpha(p_0)}{\alpha!}
\end{aligned}
$$

Comparing coefficients we conclude.

$$
\frac{GR_{H_\alpha}(\Lambda)}{\alpha!} = \sum_{\substack{\beta \in \mathbb{N}_0^n \\ |\beta| = |\alpha|}} c(P_{\Lambda_0^\perp})_\alpha^\beta \frac{H_\beta(p_0)}{\beta!}
$$

□

Example 4.1.12. Let $n = 3, k = 1$

Chapter 5

Krawtchouk Polynomials and a Discrete GRT

Here we propose a sort of "discrete Gaussian Radon transform" defined via this results of the Chapter 4.

5.1 Krawtchouk polynomials

The formula (4.1.1) gives us a path toward defining a discrete analogue to the Gaussian Radon transform. We will begin introducing the centered binomial density and corresponding Krawtchouk polynomials by way of a classic model in probability.

Consider flipping a loaded coin with probability $p \in [0,1]$ of landing on heads, and thus probability $(1-p)$ of landing on tails. The probability of k heads in N coin flips is given by the binomial probability density,

$$\binom{N}{k} p^k (1-p)^{N-k}$$

where $N \in \mathbb{N}_0$ and $k = 0, 1, \ldots, N$. It is well known that the binomial distribution can be seen as a discrete analog to the Gaussian distribution. Thus we may expect that a discrete analog to the GRT may be defined via the binomial distribution.

In order to draw a clearer line between the binomial and Gaussian densities we make use of an alternate probabilistic model: the Bernoulli random walk. Picture a particle "walking" along the integer lattice $\mathbb{Z}$, beginning at the origin $x = 0$. At every step, we flip the loaded coin described above, stepping to the right if it lands on heads, and the left if tails. After N flips, the particle's position is

$$x = (\#\text{ of heads}) - (\#\text{ of tails}) = k - (N-k) = 2k - N$$

Now the set of possible positions is $x \in [[N]] := \{-N, -N+2, \ldots, N\}$, and the probability of each is precisely the probability of landing $k = \frac{N+x}{2}$ heads, that is,

$$\binom{N}{\frac{N+x}{2}} p^{\frac{N+x}{2}} (1-p)^{N-\frac{N+x}{2}} = \binom{N}{\frac{N+x}{2}} p^{\frac{N+x}{2}} (1-p)^{\frac{N-x}{2}}$$

We will define **centered binomial density** supported on [[N]] by

$$B_N(x) = \binom{N}{\frac{N+x}{2}} p^{\frac{N+x}{2}}(1-p)^{\frac{N-x}{2}}.$$

Since the expected displacement at each (independent) step is $p-(1-p) = 2p-1$, we calculate the expected position after N steps to be

$$E(B_N) = 2Np - N.$$

Furthermore the variance of each displacement step is $4p(1-p)$ so that the variance of B_N is

$$Var(B_N) = 4Np(1-p)$$

The earlier claim that the binomial density is a discrete analog to the Gaussian density can be made explicit by the following limit relation. Let

$$z = \frac{x - E(B_N)}{\sqrt{Var(B_N)}} = \frac{x - 2Np + N}{2\sqrt{Np(1-p)}}$$

so that z has mean 0 and variance 1 with respect to B_n. Equivalently we may set

$$x = z\sqrt{Var(B_N)} + E(B_N) = 2z\sqrt{Np(1-p)} + 2Np + N$$

Then we have

$$\lim_{N\to\infty} B_N(x) = \frac{1}{(\sqrt{2\pi})} e^{-\frac{z^2}{2}} = w(z).$$

This limit is perhaps one of the simplest cases of the robust central limit theorem, but can be proven in a number of more elementary ways [**CITATION NEEDED**].

Remark 5.1.1. Under an analogous limit a random walk process x becomes a Brownian motion process z. [**CITATION NEEDED**]

Having justified the the centered binomial densities $B_N(x)$ as discrete analogs of the Gaussian $w(z)$, we would like to define a sequence of polynomials orthogonal with respect to $B_N(x)$, as the Hermite polynomials are to $w(z)$. For convenience we start with a generating function

Definition 5.1.2. The Krawtchouk polynomials $K_k^{N,p}(x)$ may be defined by the generating function

$$\sum_{k=0}^{N} K_k^{N,p}(x) y^k = (1+y)^{\frac{N+x}{2}}(1-y)^{\frac{N-x}{2}}$$

The polynomials $K_k^{N,p}(x)$ satisfy the orthogonality property

$$\sum_{x\in[[N]]} K_k^{N,p}(x)K_\ell^{N,p}(x)B_N(x) = 0, \qquad k \neq \ell$$

and can be written explicitly as

$$K_k^{N,p}(x) = \binom{N}{k} {}_2F_1\left[\begin{matrix} -k \ \ (x-N)/2 \\ -N \end{matrix}; 2\right]$$

The Krawtchouk and Hermite polynomials are related by the same limiting process.

Proposition 5.1.3. With $x = 2z\sqrt{Np(1-p)} + 2Np + N$ we have the following limit relation:

$$\lim_{N\to\infty} K_k^{N,p}(x) = H_k(z)$$

Proof. **[PROOF NEEDED]** $\square$

Now heuristically we expect a formula analagous to (eq:GRH 4.1.1) in the from

$$GR_{K_\alpha^{N,p}}(\omega, p) = K_k^{N,p}(p)\omega^\alpha, \qquad |\alpha| = k$$

where

$$K_\alpha^{N,p}(x) = \prod_{i=1}^{n} K_{\alpha_i}^{N,p}(x_i)$$

We should resolve the conflicting variables both named p.

The desired formula would represent a discrete analog for the Gaussian Radon transform on the finite cubic lattice $[[N]]^n$, however it is not yet clear how to interpret this. An immediate concern is that the cubic lattice does not admit "hyperplanes" for arbitrary p and ω. Indeed in the worst case, if ω is not rational (that is the ratios between coordinates of ω are not rational, or equivalently the hyperspherical coordinates of ω include an angle which is an irrational multiple of π) then a hyperplane orthogonal can contain at most one lattice point. In the case when ω is rational — in the sense above — we may restrict p to the projection of the lattice onto the span of ω, and sum over the lattice hyperplane $\langle x, \omega\rangle = p$ with respect to the binomial weight centered on $p\omega$. It remains to be seen whether such an interpretation agrees with the heuristic formula above.

Take the symmetric random walk, where the coin is fair ($p = 1 - p = \frac{1}{2}$). The position after N flips has distribution

$$B_N(x) = \frac{1}{2^N}\binom{N}{\frac{N+x}{2}}, \qquad x \in [[N]]$$

with expected value $E(B_N) = 0$ and variance $Var(B_N) = N$. The limiting process is then given by $B_N(x) \to w(z)$ where

$$z = \frac{x}{\sqrt{N}}$$

and further $K_k^{N,\frac{1}{2}}(x) \to H_k(z)$.

Now consider the following scenario in two dimensions. A particle starting at the origin $x = (x_1, x_2) = 0$ flips two fair coins, stepping down/up depending on the first coin, left/right depending on the second. This means the particle moves one unit up and right, up and left, down and right, or down and left, with equal probability.

Chapter 3

*

[1] L. C. Petersen, **On the relation between the multidimensional moment problem and the one-dimensional moment problem**, Math. Scand.,

[2] David Applebaum, **Lévy Processes and Stochastic Calculus**, Cambridge Studies in Advanced Mathematics, ISBN 0-521-83263-2, Cambridge University Press, 2004

[3] Richard F. Bass, **Probabilistic Techniques in Analysis**, Probability and its Applications, ISBN 0-387-94387-0, Springer Verlag, 1995

[4] Richard F. Bass, **Diffusions and Elliptic Operators**, Probability and its Applications, ISBN 0-387-98315-5, Springer Verlag, 1998

[5] Richard F. Bass and Takashi Kumagai, **Symmetric Markov Chains on $\mathbb{Z}^d$ with Unbounded Range**, Transactions of the American Mathematical Society, Vol. 30, Nr 4, 2041-2075 (2008), AMS, 2008

[6] Zhen-Qing Chen, Zhongmin Qian, Yaozhong Hu and Weian Zheng, **Stability and Approximations of Symmetric Diffusion Semigroups and Kernels**, Journal of Functional Analysis, 152, 225-280 (1998), Academic Press, 1998

[7] Rama Cont and Peter Tankov, **Financial modelling with jump processes**, Financial Mathematics Series, ISBN 1-5848-8413-4, Chapman & Hall/CRC, 2004

[8] A. De Masi, P.A. Ferrari, S. Goldstein, W.D. Wick, **An Invariance Principle for Reversible Markov Processes**, Journal of Statistical Physics, Vol. 55, 787-855 (1989), Springer Verlag, 1989

[9] Joseph L. Doob, **Classical Potential Theory and Its Probabilistic Counterpart**, Classics in Mathematics ISBN 3-540-41206-9, Springer Verlag, 1984

[10] Richard Durrett, **Probability Theory and Examples**, Wadsworth & BrooksCole StatisticsProbability Series, ISBN 0-534-13206-5, BrooksCole Publishing Company, 1991

[11] Richard Durrett, **Stochastic Calculus**, Probability and Stochastics Series, ISBN 0-8493-8071-5, CRC Press, 1996

[12] Lawrence C. Evans, **Partial Differential Equations**, Graduate Studies in Mathematics, ISBN 0-821-80772-2, American Mathematical Society, 1998

[13] William Feller, **An Introduction to Probability Theory and Its Applications Vol. II** (2nd Ed.), ISBN 0-471-25709-5, John Wiley & Sons Inc., 1990

[14] G.H. Hardy, J.E. Littlewood and G. Polya, **Inequalities** (2nd Ed.), ISBN 0-521-05206-8, Cambridge University Press, 1952

[15] Warren P. Johnson, **The Curious History of Faà di Bruno's Formula**, American Mathematical Monthly, Vol. 109, 217-234, (March 2003), MAA, 2003

[16] Ioannis Karatzas and Steven Schreve, **Brownian Motion and Stochastic Calculus** (2nd Ed.), ISBN 0-387-97655-8, Springer Verlag, 1988, 1991

[17] Hiroshi Kunita, **Stochastic Flows and Stochastic Differential Equations**, ISBN 0-521-59925-3, Cambridge University Press, 1990, 1997

[18] Jean-Pierre Lepeltier and Bernard Marchal, **Problème des martingales et équations différentielles stochastiques associées à un opérateur inégro-différentiel**, Annales de l' Institut Henri Poincaré (B) Vol. 12, 43-103, (1976)

[19] Zhi-Ming Ma and Michael Röckner, **Introduction to the Theory of (Non-symmetric) Dirichlet Forms**, ISBN 3-540-55848-9, Springer Verlag, 1992

[20] Bernt Øksendal, **Stochastic Differential Equations** (6th Ed.), ISBN 3-540-04758-1, Springer Verlag, 1985, 2003

[21] L.C.G. Rogers and D. Williams, **Diffusions, Markov Processes and Martingales, Vol. II : Itô Calculus** (2nd Ed.), ISBN 0-521-77593-0, Cambridge University Press, 2000

[22] Walter Rudin, **Functional Analysis**, ISBN 0-070-54236-8, McGraw-Hill Science and Engineering, 1991

[23] A.V. Skorokhod, **Studies in the Theory of Random Processes**, ISBN 9-780-48664240-6, Reading, Mass., Addison-Wesley Pub. Co., 1965

[24] D.W. Stroock and W. Zheng, **Markov Chain Approximations to Symmetric Diffusions**, Annales de l'Institut Henri Poincaré (B), Vol. 33, 619-649, (1997)