Lecture 3

Harmonic Analysis

Recall that if $f \in L^2([0\,,2\pi]^n)$, then we can write f as

$$f(x) = \sum_{k\in\mathbb{Z}^n} \mathrm{e}^{ik\cdot x}\hat{f}_k, \tag{3.1}$$

F:L

where the convergence takes place in $L^2([0\,,2\pi]^n)$ and $\hat{f}_k$ is the "kth Fourier coefficient" of f; that is,

$$\hat{f}_k := (2\pi)^{-n}\int_{[0,2\pi]^n} \mathrm{e}^{ik\cdot x} f(x)\,\mathrm{d}x \qquad \text{for all } k \in \mathbb{Z}^n.$$

Eq. (3.1) is the starting point of harmonic analysis in the Lebesgue space $[0\,,2\pi]^n$. In this chapter we develop a parallel theory for the Gauss space $(\mathbb{R}^n\,,\mathscr{B}(\mathbb{R}^n)\,,\mathbb{P}_n)$.

1. Hermite Polynomials in Dimension One

Before we discuss the general n-dimensional case, let us consider the special case that $n = 1$. We may observe the following elementary computations:

$$\gamma_1'(x) = -x\gamma_1(x),\ \ \gamma_1''(x) = (x^2-1)\gamma_1(x),\ \ \gamma_1'''(x) = -(x^3-3x)\gamma_1(x),\ \ldots$$

etc. It follows from these computations, and induction, that

$$\gamma_1^{(k)}(x) = (-1)^k H_k(x)\gamma_1(x) \qquad [k \geqslant 0,\ x \in \mathbb{R}],$$

where H_k is a polynomial of degree at most k. Moreover,

$$H_0(x) := 1,\ \ H_1(x) := x,\ \ H_2(x) := x^2-1,\ \ H_3(x) := x^3-3x,\ \ldots$$

etc.

Definition 1.1. H_k is called the *Hermite polynomial* of degree $k \geqslant 0$.

I should warn you that different authors normalized their Hermite polynomials differently than I have here. So my notation might differ from theirs in certain places.

lem:Hermite **Lemma 1.2.** *The following holds for all $x \in \mathbb{R}$ and $k \geqslant 0$:*

(1) $H_{k+1}(x) = xH_k(x) - H_k'(x)$;
(2) $H_{k+1}'(x) = (k+1)H_k(x)$; *and*
(3) $H_k(-x) = (-1)^k H_k(x)$.

This simple lemma teaches us a great deal about Hermite polynomials. For instance, we learn from (1) and induction that H_k is a polynomial of degree exactly k for every $k \geqslant 0$. Moreover, the coefficient of x^k in $H_k(x)$ is one for all $k \geqslant 0$; that is, $H_k(x) - x^k$ is a polynomial of degree at most $k-1$ for all $k \geqslant 1$.

Proof. We prove part (1) by direct computation:

$$\begin{aligned}(-1)^{k+1}H_{k+1}(x)\gamma_1(x) = \gamma_1^{(k+1)}(x) = \frac{\mathrm{d}}{\mathrm{d}x}\gamma_1^{(k)}(x) &= (-1)^k\frac{\mathrm{d}}{\mathrm{d}x}\left[H_k(x)\gamma_1(x)\right]\\ &= (-1)^k\left[H_k'(x)\gamma_1(x) + H_k(x)\gamma_1'(x)\right]\\ &= (-1)^k\left[H_k'(x) - xH_k(x)\right]\gamma_1(x).\end{aligned}$$

Divide both sides by $(-1)^{k+1}\gamma_1(x)$ to finish.

Part (2) is clearly correct when $k = 0$. We now apply induction: Suppose $H_{j+1}'(x) = (j+1)H_j(x)$ for all $0 \leqslant j \leqslant k-1$. We plan to prove this for $j = k$. By (1) and the induction hypothesis, $H_{k+1}(x) = xH_k(x) - kH_{k-1}(x)$. Therefore, we can differentiate to find that

$$H_{k+1}'(x) = H_k(x) + xH_k'(x) - kH_{k-1}'(x) = H_k(x) + kxH_{k-1}(x) - kH_{k-1}'(x),$$

thanks to a second appeal to the induction hypothesis. Because of (1), $xH_{k-1}(x) - H_{k-1}'(x) = H_k(x)$. This proves that $H_{k+1}'(x) = (k+1)H_k(x)$, and (2) follows.

We apply (1) and (2), and induction, in order to see that H_k is odd [and H_k' is even] if and only if k is. This proves (3). □

The following is the *raison d'être* for our study of Hermite polynomials. Specifically, it states that the sequence $\{H_k\}_{k=0}^{\infty}$ plays the same sort of harmonic-analyatic role in the 1-dimensional Gauss space $(\mathbb{R}\,, \mathscr{B}(\mathbb{R})\,, \mathbb{P}_1)$ as do the complex exponentials in Lebesgue spaces.

th:Hermite:1 **Theorem 1.3.** *The collection*

$$\left\{ \frac{1}{\sqrt{k!}} H_k \right\}_{k=0}^{\infty}$$

is a complete, orthonormal basis in $L^2(\mathbb{P}_1)$*.*

Before we prove Theorem 1.3 let us mention the following corollary.

co:Hermite:1 **Corollary 1.4.** *For every* $f \in L^2(\mathbb{P}_1)$*,*

$$f = f(Z) = \sum_{k=0}^{\infty} \frac{1}{k!} \langle f\,, H_k \rangle_{L^2(\mathbb{P}_1)} H_k(Z) = \sum_{k=0}^{\infty} \frac{1}{k!} \mathbb{E}\left[f H_k \right] H_k \qquad \textit{a.s.}$$

To prove this we merely apply Theorem 1.3 and the Riesz–Fischer theorem. Next is another corollary which also has a probabilistic flavor.

co:Hermite:Wiener:1 **Corollary 1.5** (Wiener XXX)**.** *For all* $f\,, g \in L^2(\mathbb{P}_1)$*,*

$$\mathbb{E}[fg] = \sum_{k=0}^{\infty} \frac{1}{k!}\, \mathbb{E}[f H_k]\, \mathbb{E}[g H_k], \textit{ and}$$

$$\text{Cov}(f\,, g) = \sum_{k=1}^{\infty} \frac{1}{k!}\, \mathbb{E}[f H_k]\, \mathbb{E}[g H_k].$$

Proof. Multiply both sides of the first identity of Corollary 1.4 by $g(x)$ and integrate $[\mathrm{d}\mathbb{P}_1]$ in order to obtain

$$\langle g\,, f \rangle_{L^2(\mathbb{P}_1)} = \sum_{k=0}^{\infty} \frac{1}{k!} \langle f\,, H_k \rangle_{L^2(\mathbb{P}_1)} \langle g\,, H_k \rangle_{L^2(\mathbb{P}_1)}.$$

The exchange of sums and integrals is justified by Fubini's theorem.

The preceding is another way to say the first result. The second follows from the first and the fact that $H_0 \equiv 1$. □

We now prove Theorem 1.3.

Proof of Theorem 1.3. Recall the adjoint operator A from (2.3), page 27. Presently, $n = 1$; therefore, in this case, A maps a scalar function to scalar function. Since polynomials are in the domain of definition of A [Proposition 3.3], Parts (1) and (2) of Lemma 1.2 respectively say that:[1]

$$H_{k+1} = AH_k \quad \text{and} \quad DH_{k+1} = (k+1)H_k \qquad \text{for all } k \geqslant 0. \tag{3.2}$$ A:D:H

[1] It is good to remember that H_k plays the same role in the Gauss space $(\mathbb{R}\,, \mathscr{B}(\mathbb{R})\,, \mathbb{P}_1)$ as does the monomial x^k in the Lebesgue space. Therefore, $DH_{k+1} = (k+1)H_k$ is the analogue of the statement that $\mathrm{d}(x^{k+1})/\mathrm{d}x = (k+1)x^k$. As it turns out the adjoint operator behaves a little like an integral operator and the identity $AH_k = H_{k+1}$ is the Gaussian analogue of the anti-derivative identity $\int x^k \,\mathrm{d}x \propto x^{k+1}$, valid in Lebesgue space.

Consequently,

$$\begin{aligned}\mathbb{E}(H_k^2) &= \int_{-\infty}^{\infty} H_k^2 \, \mathrm{d}\mathbb{P}_1 = \int_{-\infty}^{\infty} H_k(AH_{k-1}) \, \mathrm{d}\mathbb{P}_1 = \int_{-\infty}^{\infty} (DH_k) H_{k-1} \, \mathrm{d}\mathbb{P}_1 \\ &= k \int_{-\infty}^{\infty} H_{k-1}^2 \, \mathrm{d}\mathbb{P}_1 = k \, \mathbb{E}(H_{k-1}^2).\end{aligned}$$

Since $\mathbb{E}(H_0^2) = 1$, induction shows that $\mathbb{E}(H_k^2) = k!$ for all integers $k \geqslant 0$. Next we prove that

$$\mathbb{E}(H_k H_{k+\ell}) := \int H_k H_{k+\ell} \, \mathrm{d}\mathbb{P}_1 = 0 \qquad \text{for integers } \ell > k \geqslant 0. \tag{3.3}$$

Hermite:off:diag

By (3.2),

$$\begin{aligned}\mathbb{E}(H_k H_{k+\ell}) = \int H_k (A H_{k+\ell-1}) \, \mathrm{d}\mathbb{P}_1 &= \int (DH_k) H_{k+\ell-1} \, \mathrm{d}\mathbb{P}_1 \\ &= k \int H_{k-1} H_{k+\ell-1} \, \mathrm{d}\mathbb{P}_1.\end{aligned}$$

Now iterate this identity to find that

$$\mathbb{E}\left(H_k H_{k+\ell}\right) = k! \int H_0 H_\ell \, \mathrm{d}\mathbb{P}_1 = k! \int_{-\infty}^{\infty} H_\ell(x) \gamma_1(x) \, \mathrm{d}x = 0,$$

since $H_\ell \, \gamma_1 = (-1)^\ell \gamma_1^{(\ell)}$. It follows that $\{(k!)^{-1/2} H_k\}_{k=0}^{\infty}$ is an orthonormal sequence of elements of $L^2(\mathbb{P}_1)$.

In order to complete the proof, suppose $f \in L^2(\mathbb{P}_1)$ is orthogonal—in $L^2(\mathbb{P}_1)$—to H_k for all $k \geqslant 0$. It remains to prove that $f = 0$ $\mathbb{P}_1$-a.s.

Part (1) of Lemma 1.2 shows that $H_k(x) = x^k - p(x)$ where p is a polynomial of degree $k-1$ for every $k \geqslant 1$. Consequently, the span of $H_0, \ldots, H_k$ is the same as the span of the monomials $1, x, \cdots, x^k$ for all $k \geqslant 0$, and hence $\int_{-\infty}^{\infty} f(x) x^k \gamma_1(x) \, \mathrm{d}x = 0$ for all $k \geqslant 0$. Multiply both sides by $(-it)^k/k!$ and add over all $k \geqslant 0$ in order to see that

$$\int_{-\infty}^{\infty} f(x) \mathrm{e}^{-itx} \gamma_1(x) \, \mathrm{d}x = 0 \qquad \text{for all } t \in \mathbb{R}. \tag{3.4}$$

pre:Hermite

If the Fourier transform $\hat{g}$ of a function $g \in C_c(\mathbb{R})$ is absolutely integrable, then by the inversion theorem of Fourier transforms,

$$g(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \mathrm{e}^{-itx} \hat{g}(t) \, \mathrm{d}t \qquad \text{for all } x \in \mathbb{R}.$$

Multiply both sides of (3.4) by $\hat{g}(t)$ and integrate [dt] in order to see from Fubini's theorem that $\int fg \, \mathrm{d}\mathbb{P}_1 = 0$ for all $g \in C_c(\mathbb{R})$ such that $\hat{g} \in L^1(\mathbb{R})$. Since the class of such functions g is dense in $L^2(\mathbb{P}_1)$, it follows that $\int fg \, \mathrm{d}\mathbb{P}_1 = 0$ for every $g \in L^2(\mathbb{P}_1)$. Set $g \equiv f$ to see that $f = 0$ a.s. □

Finally, I mention one more corollary.

co:Nash **Corollary 1.6** (Nash's Poincaré Inequality). *For all* $f \in \mathbb{D}^{1,2}(\mathbb{P}_1)$,

$$\mathrm{Var}(f) \leqslant \mathbb{E}\left(|Df|^2\right).$$

Proof. We apply Corollary 1.5 and (3.2) to obtain

$$\begin{aligned}\mathrm{Var}(f) &= \sum_{k=0}^{\infty} \frac{1}{(k+1)!}\,\left|\mathbb{E}[fH_{k+1}]\right|^2 = \sum_{k=0}^{\infty} \frac{1}{(k+1)!}\,\left|\mathbb{E}[fA(H_k)]\right|^2\\ &= \sum_{k=0}^{\infty} \frac{1}{(k+1)!}\,\left|\mathbb{E}[D(f)H_k]\right|^2 \leqslant \sum_{k=0}^{\infty} \frac{1}{k!}\,\left|\mathbb{E}[D(f)H_k]\right|^2.\end{aligned}$$

The right-most quantity is equal to $\mathbb{E}(|Df|^2)$, thanks to Corollary 1.5. □

2. Hermite Polynomials in General Dimensions

For every $k \in \mathbb{Z}^n_+$ and $x \in \mathbb{R}^n$ define

$$\mathcal{H}_k(x) := \prod_{j=1}^{n} H_{k_j}(x_j) \qquad [x \in \mathbb{R}^n].$$

These are n-variable extensions of Hermite polynomials. Though, when $n = 1$, we will continue to write $H_k(x)$ in place of $\mathcal{H}_k(x)$ in order to distinguish the high-dimensional case from the case $n = 1$.

Clearly, $x \mapsto \mathcal{H}_k(x)$ is a polynomial, in n variables, of degree k_j in the variable x_j. For instance, when $n = 2$,

$$\begin{aligned}&\mathcal{H}_{(0,0)}(x) = 1, && \mathcal{H}_{(1,0)}(x) = x_1,\ \mathcal{H}_{(0,1)}(x) = x_2,\\ &\mathcal{H}_{(1,1)}(x) = x_1x_2, && \mathcal{H}_{(1,2)}(x) = x_1(x_2^2 - 1),\ \ldots\end{aligned}$$

etc. Because each measure $\mathbb{P}_n$ has the product form $\mathbb{P}_n = \mathbb{P}_1 \times \cdots \times \mathbb{P}_1$, Theorem 1.3 immediately extends to the following.

th:Hermite **Theorem 2.1.** *For every integer* $n \geqslant 1$, *the collection*

$$\left\{\frac{1}{\sqrt{k!}}\,\mathcal{H}_k\right\}_{k \in \mathbb{Z}^n_+}$$

is a complete, orthonormal basis in $L^2(\mathbb{P}_n)$, *where*

$$k! := \prod_{q=1}^{n} k_q! \qquad \textit{for all } k \in \mathbb{Z}^n_+.$$

Corollaries 1.4 has the following immediate extension as well.

co:Hermite **Corollary 2.2.** *For every $n \geqslant 1$ and $f \in L^2(\mathbb{P}_n)$,*

$$f = \sum_{k \in \mathbb{Z}_+^n} \frac{\mathbb{E}(f \mathcal{H}_k)}{k!} \mathcal{H}_k,$$

a.s., where the infinite sum converges in $L^2(\mathbb{P}_n)$.

Similarly, the following immediate extension of Corollary 1.5 computes the covariance between two arbitrary square-integrable random variables in the Gauss space.

co:Hermite:Wiener **Corollary 2.3** (Wiener XXX). *For all $n \geqslant 1$ and $f, g \in L^2(\mathbb{P}_n)$,*

$$\mathbb{E}[fg] = \sum_{k \in \mathbb{Z}_+^n} \frac{1}{k!} \mathbb{E}[f\mathcal{H}_k]\,\mathbb{E}[g\mathcal{H}_k], \text{ and}$$

$$\mathrm{Cov}(f, g) = \sum_{\substack{k \in \mathbb{Z}_+^n \\ k \neq 0}} \frac{1}{k!} \mathbb{E}(f\mathcal{H}_k)\,\mathbb{E}(g\mathcal{H}_k).$$

The following generalizes Corollary 1.6 to several dimensions.

pr:Nash **Proposition 2.4** (The Poincaré Inequality). *For all $f \in \mathbb{D}^{1,2}(\mathbb{P}_n)$,*

$$\mathrm{Var}(f) \leqslant \mathbb{E}\left(\|Df\|^2\right).$$

Proof. By Corollary 2.2, the following holds a.s. for all $1 \leqslant q \leqslant n$:

$$D_q f = \sum_{k \in \mathbb{Z}_+^n} \frac{\mathbb{E}[D_q(f)\mathcal{H}_k]}{k!} \mathcal{H}_k = \sum_{k \in \mathbb{Z}_+^n} \frac{\mathbb{E}[f A_q(\mathcal{H}_k)]}{k!} \mathcal{H}_k,$$

where we recall A_q denotes the qth coordinate of the vector-valued adjoint operator. By orthogonality and (3.2),

$$\begin{aligned}
\mathbb{E}\left(\|Df\|^2\right) &= \sum_{q=1}^n \sum_{k \in \mathbb{Z}_+^n} \frac{1}{k!} \left|\mathbb{E}[f A_q(\mathcal{H}_k)]\right|^2 \\
&= \sum_{q=1}^n \sum_{k \in \mathbb{Z}_+^n} \frac{1}{k!} \left|\mathbb{E}\left[f(Z) H_{k_q+1}(Z_q) \prod_{\substack{1 \leqslant \ell \leqslant n \\ \ell \neq q}} H_{k_\ell}(Z_\ell)\right]\right|^2.
\end{aligned}$$

Fix an integer $1 \leqslant q \leqslant n$ and relabel the inside sum as $j_\ell := k_\ell$ if $\ell \neq q$ and $j_q := k_q + 1$. In this way we find that

$$\begin{aligned}\mathbb{E}\left(\|Df\|^2\right) &\geqslant \sum_{q=1}^{n} \sum_{\substack{j \in \mathbb{Z}_+^n \\ j_q \geqslant 1}} \frac{1}{j_1! \cdots j_n!} \left| \mathbb{E}\left[f(Z) H_{j_q}(Z_q) \prod_{\substack{1 \leqslant \ell \leqslant n \\ \ell \neq q}} H_{j_\ell}(Z_\ell) \right] \right|^2 \\ &= \sum_{q=1}^{n} \sum_{\substack{j \in \mathbb{Z}_+^n \\ j_q \geqslant 1}} \frac{1}{j_1! \cdots j_n!} \left| \mathbb{E}\left[f \mathscr{H}_j \right] \right|^2.\end{aligned}$$

using only the fact that $1/(j_q - 1)! > 1/j_q!$. This completes the proof since the right-hand side is simply

$$\sum_{j \in \mathbb{Z}_+^n} \frac{1}{j_1! \cdots j_n!} \left| \mathbb{E}\left[f \mathscr{H}_j \right] \right|^2 - |\mathbb{E}[f \mathscr{H}_0]|^2,$$

which is equal to the variance of $f(Z)$ [Corollary 2.3]. □

Consider a Lipschitz-continuous function $f : \mathbb{R}^n \to \mathbb{R}$. Recall [Example 1.6, page 22] that this means that $\mathrm{Lip}(f) < \infty$, where

$$\mathrm{Lip}(f) := \sup_{\substack{x,y \in \mathbb{R}^n \\ x \neq y}} \frac{|f(x) - f(y)|}{\|x - y\|}. \tag{3.5}$$

Lip

Since $f \in \mathbb{D}^{1,2}(\mathbb{P}_n)$ and $\|Df\| \leqslant \mathrm{Lip}(f)$ a.s., Nash's Poincaré inequality has the following ready consequence:

co:Nash:Lip **Corollary 2.5.** *For every Lipschitz-continuous function $f : \mathbb{R}^n \to \mathbb{R}$,*

$$\mathrm{Var}(f) \leqslant |\mathrm{Lip}(f)|^2.$$

If f is almost constant, then $f \approx \mathbb{E}(f)$ with high probability and hence $\mathrm{Var}(f) \approx 0$. The preceding estimate is an *a priori* way of saying that "in high dimensions, most Lipschitz-continuous functions are almost constant." This assertion is somewhat substantiated by the following two examples.

Example 2.6. The function $f(x) := n^{-1} \sum_{i=1}^n x_i$ is Lipschitz continuous and $\mathrm{Lip}(f) = 1/n$. In this case, Corollary 2.5 implies that

$$\mathrm{Var}\left(n^{-1} \textstyle\sum_{i=1}^n Z_i\right) \leqslant n^{-1}.$$

The inequality is in fact an identity; this example shows that the bound in the Poincaré inequality can be attained.

Example 2.7. For a more interesting example consider either the function $f(x) := \max_{1\leqslant i\leqslant n} |x_i|$ or the function $g(x) := \max_{1\leqslant i\leqslant n} x_i$. Both f and g are Lipschitz-continuous functions with Lipschitz constant at most 1; for example,

$$|f(x) - f(y)| \leqslant \|x - y\|.$$

A similar calculation shows that $\mathrm{Lip}(g) \leqslant 1$ also. Nash's inequality implies that

$$\mathrm{Var}(M_n) \leqslant 1,^{2} \tag{3.6}$$ Var:max

where M_n denotes either $\max_{1\leqslant i\leqslant n} Z_i$ or $\max_{1\leqslant i\leqslant n} |Z_i|$. This is a nontrivial result about, for example, the absolute size of the *centered* random variable $M_n - \mathrm{E}\, M_n$. The situation changes completely once we remove the centering. Indeed by Proposition 1.3 (p. 7) and Jensen's inequality,

$$\mathbb{E}(M_n^2) \geqslant |\mathbb{E}(M_n)|^2 = (2 + o(1)) \log n \qquad \text{as } n \to \infty.$$

Similar examples can be constructed for more general Gaussian random vectors than Z, thanks to the following.

pr:Poincare:X **Proposition 2.8.** *Suppose X is distributed as $\mathrm{N}_n(0\,, Q)$ for some positive semidefinite matrix Q, and define $\sigma^2 := \max_{1\leqslant i\leqslant n} \mathrm{Var}(X_i)$. Then,*

$$\mathrm{Var}[f(X)] \leqslant \sigma^2\, \mathbb{E}\left(\|(Df)(X)\|^2\right) \qquad \textit{for every } f \in \mathbb{D}^{1,2}(\mathbb{P}_n).$$

Proof. We can write $Q = A^2$ where A is an $n \times n$ matrix [$A :=$ a square root of Q]. Define $g(x) := f(Ax)$ [$x \in \mathbb{R}^n$], and observe that: (i) X has the same distribution as AZ; and therefore (ii) $\mathrm{Var}[f(X)] = \mathrm{Var}[g(Z)] \leqslant \mathbb{E}(\|(Dg)(Z)\|^2)$ thanks to Proposition 2.4. A density argument shows that $(D_i g)(Z) = A_{i,i}(D_i f)(X)$ a.s., whence

$$\|Dg(Z)\|^2 = \sum_{i=1}^n A_{i,i}^2 |(D_i f)(X)|^2 \leqslant \max_{1\leqslant i\leqslant n} A_{i,i}^2\, \|(Df)(X)\|^2 \qquad \text{a.s.}$$

Since $A_{i,i}^2 \leqslant \sum_{j=1}^n A_{j,i}^2 = Q_{i,i} \leqslant \sigma^2$ for all $1 \leqslant i \leqslant n$, this concludes the proof. □

ex:Var:max **Example 2.9.** Suppose X has a $\mathrm{N}_n(0\,, Q)$ distribution. We can argue as we did in the preceding proof, and see that AZ and X have the same distribution where A is a square root of Q. Therefore, $\max_{i\leqslant n} X_i$ has the same distribution as $f(Z)$ where $f(x) := \max_{i\leqslant n}(Ax)_i$ for all $1 \leqslant i \leqslant n$. We saw also that f is Lipschitz continuous with $\mathrm{Lip}(f) \leqslant \max_{i\leqslant n} \mathrm{Var}(X_i)$. Therefore, Proposition 2.8 shows that for any such random vector X,

$$\mathrm{Var}\left(\max_{1\leqslant i\leqslant n} X_i\right) \leqslant \max_{1\leqslant i\leqslant n} \mathrm{Var}(X_i). \tag{3.7}$$ eq:VarM:MVar

[2] This bound is sub optimal. The optimal bound is $\mathrm{Var}(M_n) = O(1/\log n)$.

We can prove similarly that

$$\mathrm{Var}\left(\max_{1\leqslant i\leqslant n}|X_i|\right)\leqslant\max_{1\leqslant i\leqslant n}\mathrm{Var}(X_i).$$