Fourier Multipliers, Hausdorff Measure and Classification with Scattering Transform

Author: Javier Minguillón Sánchez

Universidad Autónoma de Madrid

Contents

Fourier transform and Fourier multipliers

Let $d \in \mathbb{N}$ and $f \in L^1(\mathbb{R}^d)$, we define the *Fourier transform of f* as

$$\mathfrak{F}f(\xi) = \hat{f}(\xi) = \int_{\mathbb{R}^d} f(x)e^{-2\pi i x\cdot\xi}dx. \tag{1.1}$$

Whenever $f \in L^2(\mathbb{R}^d)$, we have $\|\mathfrak{F}f\|_2 = \|f\|_2$ (Plancherel).

Definition

Let $m \in L^\infty(\mathbb{R}^d)$ and $f \in L^2(\mathbb{R}^d)$. Consider the linear operator

$$\widehat{T_m f}(\xi) = m(\xi)\hat{f}(\xi). \tag{1.2}$$

The function m is a *Fourier multiplier* and T_m is a *Fourier multiplier operator*.

Let $1 \leq p < \infty$. If there is $C > 0$ such that $\|T_m f\|_p \leq C\|f\|_p$, we can extend to $T_m : L^p \to L^p$.

Examples of multipliers

A few examples of multipliers operator are

- Translation by $\tau \in \mathbb{R}^d$ with $m_\tau(\xi) = e^{2\pi i \tau\cdot\xi}$ through

$$T_\tau f(x) = f(x-\tau) = \mathfrak{F}^{-1}\left(e^{2\pi i \tau\cdot\xi}\widehat{f}(\xi)\right)(x).$$

- Partial differentiation with respect to x_i, $i = 1, ..., d$ through

$$\frac{\partial}{\partial x_i} f(x) = \mathfrak{F}^{-1}\left((2\pi i \xi_i)\widehat{f}(\xi))\right).$$

- A low-pass filter through

$$\Phi_r f(x) = \mathfrak{F}^{-1}\left(\chi_{R_r}(\xi)\widehat{f}(\xi)\right),$$

where $R_r = [-r, r] \times \cdots \times [-r, r]$ and $\chi_.$ is the indicator function.

Properties of multipliers

Below we have a property related to the composition of a multiplier and an affine transformation.

Proposition

Fix $1 \leq p < \infty$. *Let* $T_m : L^p(\mathbb{R}^d) \to L^p(\mathbb{R}^d)$ *be a multiplier operator with multiplier* m *and* $B \in GL_d(\mathbb{R}^d)$. *Let* $\tilde{m}(\xi) = m(B\xi + \xi_0)$. *Then the operator norms coincide,*

$$\|T_{\tilde{m}}\|_{p,L^p} = \sup_{\substack{f \in L^p \\ \|f\|_p = 1}} \|T_{\tilde{m}} f\|_p = \sup_{\substack{f \in L^p \\ \|f\|_p = 1}} \|T_m f\|_p = \|T_m\|_{p,L^p}$$

.

Properties of multipliers

The following is property that holds when some (differentiable) multiplier and its derivatives have a decay.

Theorem (Hörmander-Mikhlin)

Let $m \in \mathcal{C}^k(\mathbb{R}^d \setminus \{0\})$ where $k \in \mathbb{N}$ and $k > \frac{d}{2}$. Suppose there is $C > 0$ such that, for all $\alpha = (\alpha_1, ..., \alpha_d)$ with $|\alpha| = k$,

$$\left| \left(\frac{\partial}{\partial x} \right)^{\alpha} m(x) \right| \leq C|x|^{-|\alpha|}, \tag{1.3}$$

Then, there is $A_p > 0$ such that for all $f \in L^p(\mathbb{R}^d)$,

$$\|T_m f\|_p \leq A_p \|f\|_p, \quad 1 < p < \infty. \tag{1.4}$$

The Disk Multiplier

Definition

Let $d \in \mathbb{N}$ and $f \in L^2(\mathbb{R}^d)$. The *r-disk multiplier* is the operator given by

$$\widehat{D_r f}(\xi) = \widehat{f}(\xi)\chi_{|\xi|\le r}(\xi). \tag{2.1}$$

Will $D_r f \xrightarrow[r\to\infty]{} f$ in some sense?

- In the $L^2(\mathbb{R}^d)$ norm, thanks to Plancherel's identity.
- In the L^p norm on $d = 1$, thanks to the Hilbert transform.
- Not in the L^p norm $d > 1$, $p \neq 2$. It is related to the unboundedness of the operator D_r (Fefferman).

The Counterexample

The reductions to prove that D_r is unbounded for $d > 1$, $p \neq 2$.

- For all $r > 0$, $\|D_r\|_{p,L^p} = \|D_1\|_{p,L^p}$.
- If $1/p + 1/q = 1$, then $\|D_1\|_{p,L^p} = \|D_1\|_{q,L^q}$.
- For $d \geq 2$, we have $\|D_1\|_{p,L^p(\mathbb{R}^{d-1})} \leq \|D_1\|_{p,L^p(\mathbb{R}^d)}$ (de Leeuw).

It suffices to build a counterexample to boundedness for $r = 1, p > 2$ and $d = 2$.

The counterexample we build relies on the following transformations.

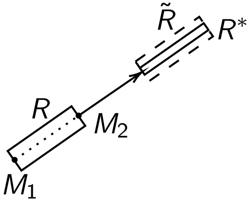

Figure: The *transformations* R^* y $\tilde{R}$ of a rectangle R.

The Counterexample

The construction relies on two lemmas.

Lemma (1)

Let $p > 2$. There exist constants $0 < C_1 \leq C_2$ such that, given $R \subset \mathbb{R}^d$ with dimensions $A^2 \times aA$ with $A, a > 0$ large enouch, the following holds.
There is a smooth function $f_R : \mathbb{R}^2 \to \mathbb{R}$ with $\operatorname{supp} f_R \subset R$ so that $|f_R| \leq 1$ in R and $C_1 \leq |D_1 f_R| \leq C_2$ in R^*.

Using Besicovitch sets, one can prove the next lemma.

Lemma (2)

Let $\epsilon > 0, a > 0$. If $A > 0$ is large enough, then there is a finite colection $\mathcal{R}$ of $A^2 \times aA$ pairwise disjoint rectangles satisfying $\left|\bigcup_{R\in\mathcal{R}} R^*\right| \leq \epsilon \left|\bigcup_{R\in\mathcal{R}} R\right|$.

Fixing ϵ and using the two lemmas, one constructs $f_\epsilon = f(x) = \sum_{R\in\mathcal{R}} \sigma_R f_R(x)$, where $\sigma_R \in \{-1, 1\}$. There is a choice of the σ_R so that $\|D_1 f\|_p^p \geq C_p \epsilon^{1-\frac{p}{2}} \|f\|_p^p$.

The Spherical Operator

Denote by $d\sigma$ the normalized measure on the unit sphere S^{d-1} inside $\mathbb{R}^d$. Take $\widehat{d\sigma}(\xi) = \int_{S^{d-1}} e^{-2\pi i x\cdot\xi} d\sigma(x)$.

Definition

Let $d \geq 2$, $t > 0$ and $f \in \mathscr{S}(\mathbb{R}^d)$. The *$t-$spherical mean of f* is

$$S_t f(x) = \int_{S^{d-1}} f(x - ty) d\sigma(y) = \mathfrak{F}^{-1}\left(\widehat{f}(\xi)\widehat{d\sigma}(t\xi)\right)(x).$$

The operator norm is $\|S_t\|_{p,L^p} \leq 1$ for all $1 \leq p \leq \infty$.

Given $1 < p < \infty$, we have that $S_t f \underset{t\to 0}{\longrightarrow} f$ in several senses:

- If $f \in L^p(\mathbb{R}^d)$, then we have $S_t f \to f$ in the L^p norm.
- Provided that $p > \frac{d}{d-1}$ as well, then we have $S_t f \to f$ a.e..

Spherical Maximal Operator and Rubio de Francia

The *spherical maximal operator* is $M_\sigma f(x) = \sup_{t>0} |S_t f(x)|$.

Theorem (Stein, Bourgain)

Given $d \geq 2$ and $p > \frac{d}{d-1}$, there is $C > 0$ such that, for all $f \in L^p(\mathbb{R}^d)$,

$$\|M_\sigma f\|_p \leq C\|f\|_p.$$

This can be proven by a theorem due to Rubio de Francia and the fact that $|\widehat{d\sigma}(\xi)| \leq C|\xi|^{\frac{d-1}{2}}$.

Theorem (Rubio de Francia)

Let $d\mu$ be a compactly supported finite Borel measure and let $\{T_t\}_{t>0}$ be a family of multiplier operators defined by

$$\widehat{T_t f}(\xi) = \widehat{f}(\xi)\widehat{d\mu}(t\xi).$$

Let $M_\mu f(x) = \sup_{t>0} |T_t f|(x)$. If there is $C > 0$ such that $|\widehat{d\mu}(\xi)| \leq C|\xi|^{-a}, \xi \in \mathbb{R}^d$, then M_μ is bounded for all $p > \frac{2a+1}{2a}$ with $a > 1/2$.

Hausdorff measure and Hausdorff dimension

Let $s > 0$, $E \subset \mathbb{R}^d$ and $\delta > 0$. We define

$$\mathcal{H}_s^\delta(E) = \inf \left\{ \sum_{j=1}^{\infty} diam(E_j)^s : E \subset \bigcup_{j=1}^{\infty} E_j,\ diam(E_j) \leq \delta \right\}.$$

As it turns out, $m_s(E) = \lim_{\delta\to 0} \mathcal{H}_s^\delta(E)$ defines a measure on the Borelian sets of $\mathbb{R}^d$: the *s-dimensional Haussdorff measure*.

The measure m_s has an important property. Given $t > 0$ and A Borelian in $\mathbb{R}^d$,

- if $m_s(A) < \infty$ and $t > s$, then $m_t(A) = 0$.
- if $m_s(A) > 0$ and $t < s$, then $m_t(A) = \infty$.

Definition

Let A be a Borelian in $\mathbb{R}^d$. The *Hausdorff dimension of A* is

$$\dim A = \inf\{t : m_t(A) = 0\}$$

Construction of fractal sets

One can construct self similar sets of arbitrary Hausdorff dimension.

Theorem

Let $S_1, ..., S_m$ *be similitudes of ratio* $0 < r < 1$. *Then, there exists exactly one non empty compact set* F *such that*

$$F = \tilde{S}(F) := S_1(F) \cup \cdots \cup S_m(F).$$

We say that m similitudes are *separated* if there exists an bounded open set $U \neq \emptyset$ such that $S_i(U) \cap S_j(U) = \emptyset$ for all $1 \leq i < j \leq m$ and $\tilde{S}(U) \subset U$.

Proposition

Let F *be a compact set such that* $F = \tilde{S}(F)$ *where* $S_1, ...S_m$ *are separated similarities with ratio* $0 < r < 1$. *Then,* $\dim F = \frac{\log m}{\log\left(\frac{1}{r}\right)}$

The Geometric Measure Theory results

Lemma (Frostman)

Let A be a Borelian and $s > 0$. We have $m_s(A) > 0$ if and only if there exists a finite Borel measure $d\mu$ that is compactly supported on A and satisfies

$$d\mu(B) \leq r^s \tag{3.1}$$

for all balls B of radius r within $\mathbb{R}^d$.
Thus, $\dim A = \sup\{s > 0 : \text{there exists } d\mu \text{ such that (3.1) holds}\}$.

Given $t \in S^{d-1} \subset \mathbb{R}^d$, the projection length is $P_t(x) = x \cdot t$.

Theorem

Let A be a Borelian in $\mathbb{R}^d$ and $s = \dim A$.

i) *If $s \leq 1$, then $\dim P_t(A) = s$ for almost all $t \in S^{d-1}$ with respect to $d\sigma$.*

ii) *If $s > 1$, then $|P_t(A)| > 0$, (in particular, $\dim P_t(A) = 1$) for almost all $t \in S^{d-1}$ with respect to $d\sigma$.*

Functional Data Statistics

It is the statistical analysis of functional data. Here the random elements that we want to predict are stochastic processes. For our purposes, we say a *stochastic process* is a random element

$$X : (\Omega, \Sigma, P) \to (L^2(T), \|\cdot\|_2),$$

where T is a compact interval of $\mathbb{R}$.

The *weak expectation* of X is given by

$$EX(t) = E[X(t)] = \int_\Omega X(\omega, t) dP(\omega), \quad t \in T.$$

The *covariance function* of X is $\gamma(t, s) = E[X(t)X(s)]$ and the *covariance operator* of X is

$$\Gamma X(t) = \int_T \gamma(t, s) X(s) ds.$$

Karhunen-Loève theorem

Now we can state the decomposition theorem.

Theorem

Let X be a stochastic process with $EX(t) = 0$ and $EX^2(t) < \infty$ for all $t \in T$. Suppose that $\gamma(s,t)$ is continuous. Let $\{e_k\}_{k=1}^{\infty}$ be the set of eigenfunctions of Γ with respective eigenvalues $\{\lambda_k\}_{k=1}^{\infty}$. Then we have a decomposition

$$X(t) = \sum_{k=1}^{\infty} Z_k e_k(t),$$

where the sum converges in the $L^2(\Omega)$ norm uniformly over T and:

(i) *The $\{e_k\}_{k=1}^{\infty} \subset L^2(T)$ form an orthonormal set.*

(ii) *The $Z_k = \int_T X(t)e_k(t)dt$ are pairwise uncorrelated random variables with $E[Z_k] = 0$ and $E[Z_k^2] = \lambda_k$, $1 \leq k \leq \infty$.*

Principal Components Analysis

We want to find the directions of larger variance of a suitable process X.

Let X with $EX^2(t) \in L^2(\Omega)$ for all $t \in T$. Take $a \in L^2(T)$ and $Y = X - EX$.

By (i) of Karhunen-Loève, $\langle a, Y \rangle = \sum_{k \in \mathbb{N}} Z_k \langle a, e_k \rangle$ and, thanks to (ii), we have that

$$V(\langle a, Y \rangle) = \sum_{k \in \mathbb{N}} \lambda_j \langle a, e_j \rangle^2 \leq \lambda_{max} \sum_{k \in \mathbb{N}} \langle a, e_j \rangle^2 = \lambda_{max} \|a\|^2_{L^2(T)}.$$

Now we discard the direction $k = max$ and repeat the process.

The principal components correspond to the largest eigenvalues.

Directional wavelets and scattering transform

Let $\psi \in L^2(\mathbb{R}^2)$ and let $\phi \in L^1 \cap L^2(\mathbb{R}^2)$ with $\|\phi\|_1 = 1$ be a low-pass filter. Fix a scale $K \in \mathbb{N}$ and denote $\Phi_K = 2^{-2K}\phi(2^{-K}x)$.

The *directional wavelets* associated to $\psi \in L^2(\mathbb{R}^2)$ are given by

$$\psi_{k,\theta}(x) = 2^{-2k}\psi(2^{-k} r_\theta x), \quad k < K,\ \theta \in \Theta,$$

where Θ is a set of angles.

Fix a function $f : T_1 \times T_2 \to \mathbb{R}$ and an order $m \in \mathbb{N}$ and consider the following vector-valued functions,

$$S_{q,K} f(x) = \left(\left| \left| \cdots \left| f * \psi_{k_1,\theta_1} \right| * \cdots \right| * \psi_{k_q,\theta_q} \right| * \phi_K \right)_{\substack{k_1 < \cdots < k_q < K \\ \theta_1,\ldots,\theta_q \in \Theta}} .$$

Definition

The *scattering transform* of f is $S_K f(x) = \left(S_{q,K} f(x)\right)_{0 \le q \le m}$.

Directional wavelets and PCA

The *scattering distance* between two pictures f and g is

$$|||S_K f - S_K g|||^2 = \int \|S_K f(x) - S_K g(x)\|^2 dx. \qquad (4.1)$$

Suppose there are L types of pictures. Then, each picture sample f can be seen as a realization of a random element F_ℓ where $1 \leq \ell \leq L$.

Our objective is to classify images looking at their scattering tranform. We need to train a classifier.

We take many images of some class ℓ and estimate the covariance function of $S_K F_\ell - E[S_K F_\ell]$. After that, we diagonalize it to find D directions of high variability that form a linear space $V_{D,\ell}$. We repeat this process for all $\ell \in \{1, ..., L\}$.

The classifier is $\hat{\ell} = \arg\min_{l \leq L} |||S_K f - E[S_K F_\ell] - P_{V_{D,\ell}}(S_K f)|||$. Here $P_{V_{D,\ell}}$ is the projection over $V_{D,\ell}$.

References I

A. Cuevas
Class notes for Advanced Course in Statistics at Universidad Autónoma de Madrid

J. Duoandikoetxea.
Fourier analysis, volume 29 of *Graduate studies in mathematics*.
American Mathematical Society, Providence R.I., 2001.

L. Grafakos.
Modern Fourier analysis, volume 250 of *Graduate texts in mathematics*.
Springer, New York, third edition, 2014.

S. Mallat.
Group Invariant Scattering, 2011 https://arxiv.org/abs/1101.2286

S. Mallat, J. Bruna
Classfication with Scattering Operators, 2011
https://arxiv.org/abs/1011.3023

P. Mattila
Fourier Analysis and Hausdorff Dimension (Cambridge Studies in Advanced Mathematics).
Cambridge University Press, (2015)

References II

F. Soria.
Five lectures on Harmonic Analysis, chapter 12 of Topics in mathematical analysis.
Edited by P.Ciatti et. al. , World Scientific, 2008.

E. M. Stein.
Harmonic analysis: Real-variable methods, orthogonality, and oscillatory integrals, volume 43 of *Princeton mathematical series*.
Princeton University Press, Princeton N.J., 1993.

T. Tao.
Bochner-Riesz summation, Fefferman's disc multiplier counterexample (class notes), 1999.