THEORY OF THE RADON TRANSFORM

The development of this section is based primarily on [3], and to a lesser extent on [1] and [2]. The alternate inversion formula comes from [4] and [1].

The Radon Transform, often called the X-ray transform for its obvious connection to X-ray imaging and its central importance in X-ray computed tomography (CT), also plays a role in MRI. Although the most common method of creating an image in MRI is to collect Fourier transform data, some machines instead collect two- or three-dimensional Radon Transform data in what is called the projection reconstruction (PR) method.

1. Definition and Basic Properties

Recall from our discussion of polar notation that S^{n-1} is the set of vectors in real n-space with norm 1; in particular, S^1 is the unit circle in the plane. For a particular $\omega \in S^1$ and $r > 0$, let $l_{r,\omega} = \{x \in \mathbb{R}^2 : x \cdot \omega = r\}$. If we choose a unit vector $\omega^\perp$ orthogonal to ω such that

$$\det \begin{pmatrix} \omega_1 & \omega_2 \\ \omega_1^\perp & \omega_2^\perp \end{pmatrix} = 1.$$

(There are two vectors orthogonal to ω in $\mathbb{R}^2$. For the purposes of simply defining the Radon transform it matters not which we use—see exercises—but for some of the theorems it is useful to consistently choose $\omega^\perp$ to point in the same direction relative to ω.) Note that this line can be parameterized by the function $l : \mathbb{R} \to \mathbb{R}^2$,

$$l(t) = r\omega + t\omega^\perp.$$

Definition 1. Let f be a bounded, piecewise continuous function with compact support in $\mathbb{R}^2$. The *Radon Transform* of f is the function $\mathcal{R}f : \mathbb{R} \times L^1$ defined by

$$\mathcal{R}f(r,\omega) = \int_{l_{r,\omega}} f = \int_{\mathbb{R}} f(r\omega + t\omega^\perp)dt.$$

Remark 2. Clearly, by linearity of the integral, the Radon transform is linear. Also, because the inner product is linear, it follows that $l_{r,\omega} = l_{-r,-\omega}$, so that the Radon transform is even. Finally, it is clear from the formula as an integral that if $f \geq 0$, then $\mathcal{R}f \geq 0$. The support properties of $\mathcal{R}f$ follow from the definition: if $f(x) = 0$ for $|x| > R$, then $\mathcal{R}f(r,\omega)$ is zero whenever $|r| > R$.

Thus far, we have confined our definition of the Radon transform to bounded, piecewise continuous functions. In fact, it can be defined for all L^1 functions, with the caveat that it need only be defined for almost every (r,ω) pair.

In order to simplify the arguments below, we will associate with each $\omega \in S^1$ a (unique) angle $\theta \in [0, 2\pi)$ that ω makes with the positive x axis, and write $\theta(\omega)$ and $\omega(\theta)$ when convenient. We will also assume (with some justification from the geometric definition of a radian) that with this identification, the measure σ on S^1 is equivalent to ordinary Lebesgue measure on $[0, 2\pi)$.

Our goal is to show that $\mathcal{R}$ is a continuous linear map on L^1; in the process we show that it is well-defined for each L^1 function, except possibly on a set of measure zero. In addition, for continuity, there are two sensible norms that we could use on $S^1 \times \mathbb{R}$. The first is the ordinary L^1 norm,

$$\|\mathcal{R}f\|_1 = \int_{S^1}\int_{\mathbb{R}} |\mathcal{R}f(r,\omega)|\, dr d\sigma(\omega) = \int_0^{2\pi}\int_{\mathbb{R}} |\mathcal{R}f(r,\omega(\theta))|\, dr d\theta.$$

The second is a mix of the L^1 norm and the sup norm,

$$\|\mathcal{R}f\|_{1,u} = \sup_{\omega\in S^1}\int_{\mathbb{R}} |\mathcal{R}f(r,\omega)|\, dr.$$

It turns out that the Radon transform is continuous using either norm on the range:

Proposition 3. *The Radon transform defines a continuous linear map from $L^1(\mathbb{R}^2)$ to $L^1(S \times \mathbb{R})$, under both the norms $\|\cdot\|_1$ and $\|\cdot\|_{i,u}$.*

Proof. Let $T(\theta, s, t) = s\omega(\theta) + t\omega^\perp(\theta)$, which for any given θ is a rotation, rotating a point $(s,t) \in \mathbb{R}^2$ counterclockwise around the origin by an angle of θ. To be thorough, we check that the s, t Jacobian matrix of $T(\theta, ., .)$ is

$$D_{s,t}T(\theta, s, t) = \begin{pmatrix} \omega_1(\theta) & \omega_2(\theta) \\ \omega_1^\perp(\theta) & \omega_2^\perp(\theta) \end{pmatrix},$$

---

[0] preliminary version

which is by definition an orthogonal matrix, confirming that T is a rotation. Then, for a fixed θ, $f\circ T$ is measurable and $\|f\circ T(\theta,.,.)\|_1 = \|f\|_1$ by the change of variables formula. Since, by definition,

$$\mathcal{R}f(r,\omega) = \int_{\mathbb{R}} f\circ T(\theta(\omega),r,s)ds,$$

it follows that for each ω, the Radon transform is well-defined for almost every s. It also follows from monotonicity of the integral that

$$\begin{aligned}\int_{\mathbb{R}} |\mathcal{R}f(r,\omega)|\,dr &= \int_{\mathbb{R}}\left|\int_{\mathbb{R}} f\circ T(\theta(\omega),r,s)ds\right| dr\\ &\le \int_{\mathbb{R}}\left|\int_{\mathbb{R}} f\circ T(\theta(\omega),r,s)\right| dsdr\\ &= \|f\circ T(\theta,.,.)\|_1 = \|f\|_1\,.\end{aligned}$$

Taking the supremum, we get that $\mathcal{R}$ is bounded if we use $\|\cdot\|_{1,u}$ on the range. Boundedness under the ordinary 1-norm follows easily, since $\sigma(S^1)=2\pi$, so that

$$\|\mathcal{R}f\|_1 \le 2\pi\,\|\mathcal{R}f\|_{1,u}\,.$$

□

Recall that the convolution of two measurable functions f and g on $\mathbb{R}^n$ is defined by

$$f*g(x) = \int_{\mathbb{R}^n} f(x-y)g(y)dy,$$

assuming that the integral converges for almost every x.

Proposition 4. *Let $f,g\in L^1(\mathbb{R})$ (recall that this ensures that $f*g$ is well-defined). Then*

$$\mathcal{R}(f*g) = \mathcal{R}f *_{aff} \mathcal{R}g,$$

where $*_{aff}$ indicates convolution in the affine parameter, that is,

$$\mathcal{R}f *_{aff} \mathcal{R}g(r,\theta) = \int_{\mathbb{R}} \mathcal{R}f(r-t,\theta)\mathcal{R}g(t,\theta)dt.$$

Proof. The proof uses Fubini's theorem to change the order of integration and invariance of integrals under translations. In the fourth line, we make a change of variables $t' = t - y\cdot\omega^{\perp}$ in order to move $y\cdot\omega$ into the coefficient of ω in the argument. In the sixth line, we use the orthonormal change of variables from $\mathbb{R}^2$ to $\mathbb{R}\times\mathbb{R}$ given $y = s\omega + t\omega^perp$, and the fact that when we make this change, $y\cdot\omega$ becomes s. Let $\omega=\omega(\theta)$. Then,

$$\begin{aligned}\mathcal{R}(f*g)(r,\theta) &= \int_{\mathbb{R}} f*g(r\omega+t\omega^{\perp})dt\\ &= \int_{\mathbb{R}}\int_{\mathbb{R}^2} f(r\omega+t\omega^{\perp}-y)g(y)dydt\\ &= \int_{\mathbb{R}^2}\int_{\mathbb{R}} f(r\omega+t\omega^{\perp}-y)g(y)dtdy\\ &= \int_{\mathbb{R}^2} g(y)\int_{\mathbb{R}} f((r-y\cdot\omega)\omega+t'\omega^{\perp})dt'dy\\ &= \int_{\mathbb{R}^2} g(y)\mathcal{R}f((r-y\cdot\omega),\theta)dy\\ &= \int_{\mathbb{R}}\int_{\mathbb{R}} g(s\omega+u\omega^{\perp})\mathcal{R}f(r-s,\theta)duds\\ &= \int_{\mathbb{R}} \mathcal{R}g(s,\theta)\mathcal{R}f(r-s,\theta)ds\\ &= \mathcal{R}f *_{aff} \mathcal{R}g(r,\theta).\end{aligned}$$

□

The following fundamental fact about the Radon transform comes from the fact that, due to the fact that the integration defining $\mathcal{R}f$ is along a line orthogonal to ω, any function of the inner product $x\cdot\omega$ remains constant in the integration. The motivation for this theorem comes from the fact that the Fourier transform is defined as an integration against a bounded function

Theorem 5. *(*General Projection Slice Theorem*) Let $h : \mathbb{R} \to \mathbb{R}$ be any bounded measurable function. Then for any $x = r\omega \in \mathbb{R}^2$*

$$\int_{\mathbb{R}^2} f(x)h(r\omega \cdot x)dx = \int_{\mathbb{R}} \mathcal{R}f(t,\omega)h(rt)dt.$$

Proof. Fixing $\theta = \theta(\omega)$, let $T = T(\theta, ., .)$ be as in Proposition 3. Then, by the definition of T, it follows that

$$h(T(x_1, x_2) \cdot \omega) = x_1.$$

Using the orthogonal change of variables given by T, that is, $T(x) = t\omega + s\omega^{\perp}$, along with Fubini's theorem, we have

$$\begin{aligned} \int_{\mathbb{R}^2} f(x)h(r\omega \cdot x)dx &= \int_{\mathbb{R}^2} f \circ T(x)h(r\omega \cdot T(x))dx \\ &= \int_{\mathbb{R}} \int_{\mathbb{R}} f(t\omega + s\omega^{\perp})h(rt)dsdt \\ &= \int_{\mathbb{R}} \mathcal{R}f(t,\omega)h(rt)dt, \end{aligned}$$

where in the third line, we used the fact that $r\omega \cdot (t\omega + s\omega^{\perp}) = rt$. ☐

By the comments made before the theorem, we get the following corollary:

Corollary 6. *(*Fourier Slice Theorem*) If we let $\hat{f}$ denote the 2-dimensional Fourier transform of f and $\widetilde{\mathcal{R}f}$ denote the 1-dimensional Fourier transform of $\mathcal{R}f$ in the affine (radial) direction, then for each $(r,\omega) \in \mathbb{R} \times S^1$,*

$$\hat{f}(r\omega) = \widetilde{\mathcal{R}f}(r,\omega).$$

Since the Fourier transform of an L^1 function is unique, it follows that if $f, g \in L^1$ and $\mathcal{R}f = \mathcal{R}g$, then $f = g$ (as always, up to a set of measure zero), and hence, the Radon transform is injective.

Combining this with the Plancherel formula for the Fourier transform, and a change of variables to polar coordinates, we also get a Plancherel formula for the Radon transform: if f is both integrable and square-integrable, then

$$\int_{S^1} \int_{\mathbb{R}} \left| \widetilde{\mathcal{R}f} \right|^2 |r| \, dr d\sigma\omega = 4\pi \left\| f \right\|_2.$$

Interestingly, this shows that we *cannot* extend the Radon transform to a continuous linear map from $L^2(\mathbb{R}^2)$ to $L^2(\mathbb{R} \times S^1)$, because the presence of $|r|$ in the integral at left leads to the operator not being bounded. See Epstein for more discussion.

2. Backprojection and Inverting the Radon Transform

The Fourier slice theorem leads directly to one formula for inverting $\mathcal{R}$:

$$f = \mathcal{F}_{\mathbb{R}^2}^{-1} \mathcal{F}_{\mathbb{R}} \mathcal{R}f.$$

That is,

$$f(x) = \frac{1}{2} \int_{S^1} \int_{\mathbb{R}} \widehat{\mathcal{R}f}(t,\omega) \, |r| \, e^{2\pi i r\omega \cdot x} dr d\sigma(\omega), \tag{1}$$

where the $\frac{1}{2}$ comes from the fact that we are integrating on all of $\mathbb{R}$ in the affine parameter, rather than just $[0, \infty)$. In some settings, this leads to satisfactory inversion algorithms with actual data; methods based on this formula are sometimes used in MRI. But in X-ray CT, the inversion algorithms are usually based on an alternative inversion formula that makes use of the backprojection operator. The two formula are equivalent in an analytic sense, but they lead to different approximate inverses when applied to finite data, and the backprojection method generally gives superior results in CT.

Definition 7. We define the *backprojection operator* $\mathcal{R}^* : L^1(\mathbb{R} \times S^1) \to L^1(\mathbb{R})$ by

$$\mathcal{R}^* g(x) = \int_{S^1} g(x \cdot \omega, \omega) d\sigma(\omega).$$

Note that the backprojection operator $g(x)$ is just the average of g on all lines that intersect x. In some sense, we might expect such an operator, which offers a very approximate inverse to the Radon transform (see Epstein for a graphical example), to be involved in the inversion formula.

The notation $\mathcal{R}^*$ comes from the standard notation for the adjoint, or dual, of a linear operator.

Definition 8. Let A be a linear operator $X \to Y$ for inner product spaces X and Y. A linear operator $A^* : Y \to X$ is called the *adjoint* or *dual* of A if, for each $x \in X$ and $y \in Y$,

$$\langle Ax, y\rangle_Y = \langle x, A^* y\rangle_X .$$

Because the term dual has a different meaning in functional analysis, we will use the term adjoint here. It is worth pointing out that, from elementary linear algebra, if A is an operator between finite-dimensional spaces, and hence, representable by a matrix, then its adjoint is represented by that matrix's conjugate transpose, which is also denoted A^*.

Proposition 9. *If we restrict the Radon transform to the space of square-integrable functions so that the L^2 inner product is defined on the domain, then the restricted range is also square-integrable, and $\mathcal{R}^*$ is the adjoint to $\mathcal{R}$.*

Proof. Let $f \in L^1(\mathbb{R}^2) \cap L^2(\mathbb{R}^2)$, $g \in L^1(\mathbb{R} \times S^1) \cap L^2(\mathbb{R} \times S^1)$. Using the definitions and using Fubini's theorem to interchange the order if integration, we find that

$$\begin{aligned} \langle \mathcal{R}f, g\rangle_{L^2(\mathbb{R}\times S^1)} &= \int_{S^1}\int_{\mathbb{R}} \mathcal{R}f(r,\omega) g(r,\omega) dr d\omega \\ &= \int_{S^1}\int_{\mathbb{R}^2} f(r\omega + s\omega^\perp) g(r,\omega) ds dr d\omega \\ &= \int_{\mathbb{R}^2}\int_{S^1} f(r\omega + s\omega^\perp) g(r,\omega) d\omega ds dr \end{aligned}$$

We now let $s = r\omega + s\omega^\perp$, so that $r = x \cdot \omega$, to get

$$\begin{aligned} \langle \mathcal{R}f, g\rangle_{L^2(\mathbb{R}\times S^1)} &= \int_{\mathbb{R}^2}\int_{S^1} f(s) g(x\cdot\omega, \omega) d\omega dx \\ &= \langle f, \mathcal{R}^*(g)\rangle_{L^2(\mathbb{R}^2)} . \end{aligned}$$

□

Note that, if we insert parentheses around the inner integral in equation (1) to get

$$f(x) = \frac{1}{2}\int_{S^1} \left[\int_{\mathbb{R}} \widehat{\mathcal{R}f}(x\cdot\omega, \omega) \left|r\right| dr e^{2\pi i r \omega\cdot x} \right] d\sigma(\omega), \tag{2}$$

Then the outer integral has the same form as the backprojection operator, so we can rewrite the inversion formula as

$$f = \mathcal{R}^* G \mathcal{R} f,$$

where

$$G\mathcal{R}f(r,\omega) = \frac{1}{2}\int_{\mathbb{R}} \widehat{\mathcal{R}f}(t,\omega) e^{2\pi i r t} \left|t\right| dt$$

is called the *filtered* Radon transform. Approximate inverses based on this formula have two advantages over Fourier reconstruction methods.

This first is that the fast Fourier transform can be used only on data from uniform rectilinear samples of the Fourier transform, and in finite settings Radon transform data will lead to samples of radial slices of the Fourier transform—that is, lines through the origin—which means that the data have to be interpolated to a grid, in a process inventively named 'gridding,' before they can be used to reconstruct the objective function. Approximations to the filtered back projection formula never call for a two dimensional Fourier transform, so they avoid this issue.

The second is that, unlike in Fourier reconstruction, the data from each view can be mostly processed while the measurements from other views are still underway, which speeds up the reconstruction. This is because the filter only acts in r, so each view can be filtered independently of the others. The only step that requires data from all views is the back-projection step, which is relatively quick. The filter is shift-invariant, so it can be approximated using a discrete convolution, which, via the fast Fourier transform method, is relatively efficient.

3. Radon Transform in Higher Dimensions and an Alternative Inversion Formula

We can also define the Radon transform in higher dimensions, as a map from $L^1(\mathbb{R}^n) \to L^1(\mathbb{R} \times S^{n-1})$. The proofs of results in this section are left to the reader.

Definition 10. For $r \in \mathbb{R}$ and $\omega \in S^{n-1}$, let $l_{r,\omega}$ be the $n-1$ dimensional hyperplane

$$\{x \in \mathbb{R}^n : x \cdot \omega = r\}.$$

Example 11. In $\mathbb{R}^3$, $l_{r,\omega}$ plane perpendicular to $\vec{\omega}$ located r units from the origin. In dimensions higher than three, it is more difficult to visualize, but is still the hyperplane orthogonal to ω located r units from the origin.

Definition 12. The n-dimensional *Radon transform* $L^1(\mathbb{R}^n) \to L^1(\mathbb{R} \times S^{n-1})$ is defined, for $f \in L^1(\mathbb{R}^n)$, by

$$\mathcal{R}f(r,\omega) = \int_{l_{r,\omega}} f(x)dx,$$

where we use a measure on the hyperplane that is equivalent to Lebesgue measure on $\mathbb{R}^{n-1}$.[1]

Definition 13. The higher-dimensional *Radon adjoint,* that is, the higher-dimensional backprojectiton operator, which maps functions on $\mathbb{R} \times S^{d-1}$ to functions on R^d, is defined for $h \in L^1(\mathbb{R} \times S^{d-1}) \cap L^2(\mathbb{R} \times S^{d-1})$ by

$$\mathcal{R}^*(x) = \int\limits_{S^{d-1}} h(x \cdot \omega, \omega) d\sigma(\omega).$$

Remark 14. The formula above is equivalent to the following, more explicit formula: if $\omega \in S^{n-1}$, then choose $\{\omega_j^\perp : j = 2,..,n\}$ so that the $\omega_j^\perp$'s and ω form an orthonormal basis for $\mathbb{R}^n$. For $r \in \mathbb{R}$, we can write

$$\mathcal{R}f(r,\omega) = \int_{\mathbb{R}^n - 1} f(r\omega + x_2\omega_1^\perp + ... + x_n\omega)dx_2...dx_n.$$

We leave it to the reader to verify that the projection-slice theorem continues to hold for the n dimensional Radon transform that the back-projection operator, this time with the integral over S^{n-1}, is still its adjoint.

The case of $\mathbb{R}^3$ allows for a convenient alternative inversion formula, not based on the Fourier inversion method we used to develop filtered backprojection. This inversion formula makes use of the Laplace operator, or Laplacian, a differential operator that plays a role in the heat and wave equations.

Definition 15. Suppose $f : \mathbb{R}^d \to \mathbb{R}$. We define the *Laplace operator* of f, denoted Δf, by

$$\Delta f(\vec{x}) = \frac{\partial^2 f}{\partial x_1^2} + \frac{\partial^2 f}{\partial x_2^2} + ... + \frac{\partial^2 f}{\partial x_d^2}.$$

Proposition 16. *Let $f \in \mathcal{S}(\mathbb{R}^3)$. Then*

$$f = -\frac{1}{8\pi}\Delta(\mathcal{R}^*\mathcal{R}f). \tag{3}$$

We leave the proof to the exercises. The formula holds in a distributional sense if f is not twice differentiable, but in the exercises we will only ask the reader to prove it when f is Schwarz.

We can extend this method of inverting the Radon transform by defining positive powers of the Laplace operator:

Definition 17. We define $(-\Delta)^a$ on $\mathcal{S}'(\mathbb{R}^d)$ by

$$(-\Delta)^a f = ((2\pi\,|\xi|)^{2a}\hat{f})\check{}.$$

(Here ˇ and ˆ indicate the unitary Fourier transform and its inverse. The 2π does not appear if we use the angular Fourier transform.)

We leave it to the reader (see assignments) to verify that, when s is a positive integer, so that Δ^s would already have been defined as a composition of Δ with itself, the two definitions are equivalent.

Equipped with this definition of Δ^s, we can prove an inversion formula analogous to Equation 3 for the n-dimensional Radon transform, whose proof is again left to the exercises:

Proposition 18. *If $f \in \mathcal{S}(\mathbb{R}^n)$, then*

$$f = \frac{2}{(2\pi)^{1-n}}(-\Delta)^{(n-1)/2}(\mathcal{R}^*\mathcal{R}f)$$

References

[1] Charles Epstein. *Introduction to the Mathematics of Medical Imaging.* Prentice Hall, Upper Saddle River, 2nd edition, 2007.

[2] Eric Todd Quinto. An introduction to z-ray tomography and radon transforms. In Gestur Ólafsson and Eric Todd Quinto, editors, *Proceedings of Symposia in Applied Mathematics: American Mathematical Society Short Course, January 3-4, 2005, Atlanta, Georgia*, volume 63, pages 1–24, Providence, Rhode Island, September 2006. American Mathematical Society.

[3] A.G. Ramm and A.I. Katsevich. *The Radon Transform and Local Tomography.* CRC Press, New York, 1996.

[4] Elias Stein and Rami Shakarchi. *Fourier Analysis: an Introduction.* Princeton University Press, Princeton, 2003.

---

[1] since Lebesgue measure is invariant to translations and rotations, there is no ambiguity; we can treat any point in $l_{r,\omega}$ as the origin—it is conventional to use the point $r\omega$, which our more explicit formula implicitly does—and choose orthogonal axis any way we wish.