The Radon transform

Chris Stolk

December 18, 2014

1 Introduction

In two dimensions the Radon transform is an integral transform that maps a function to its integrals over lines. Let $\theta \in S^1$ and $s \in \mathbb{R}$ then the equation $x \cdot \theta = s$ describes a line. The Radon transform Rf of a function f in $\mathscr{S}(\mathbb{R}^2)$ is defined by

$$Rf(\theta, s) = \int_{x \cdot \theta = s} f(x)\, dx. \tag{1}$$

(Note that the improper integral converges.) Here the integral denotes the standard line integral from vector calculus. In higher dimensions the Radon transform maps a function f to its integrals over hyperplanes. It is defined by the same formula (1), but now $\theta \in S^{n-1}$, and the integration domain $\theta \cdot x = s$ is a hyperplane. The Radon transformed is named after the Austrian mathematician Johann Radon, who studied it in a paper that appeared in 1917. In particular he proved several inversion formulas in his paper.

The two-dimensional Radon transform has application in medical imaging, in particular in X-ray computed tomography. In an X-ray scanner, a beam of X-ray radiation is generated that passes through an object, see Figure 1. At the other side of the object a line of detectors is located that measures the intensity I of the rays that have passed through the object. The beam and the detectors can be rotated, such that the intensity is measured for all lines passing through the object, hence $I = I(\theta, s)$. Depending on the material inside the object, some of the X-ray radiation is absorbed. When an X-ray travels over a small distance Δx in a medium with absorbtion coefficient $f(x)$, the intensity change is $\Delta I = -f(x) I \Delta x$. This leads to a differential equation, if $x(t)$ is a parametrization of a line, with $\left| dx/dt \right| = 1$, the equation is $\frac{dI}{dt} = -f(x(t)) I$ with solution

$$I(t) = e^{-\int^t f(x(s)) ds} I_0.$$

Hence for the rays that have passed through the object we have

$$-\log(\frac{I(\theta, s)}{I_0}) = \int_{x \cdot \theta = s} f(x) = s\, dx.$$

So the reconstruction problem for f amount to the inversion of the Radon transform. For more on this problem see [6, 3]. (A 3-D scan can be done by performing a series of 2-D scans for thin layers of the object.)

As Epstein [3] notes, there are a number of other imaging methods in medical imaging that involve inverting certain integral transforms. The methods used in studying the Radon

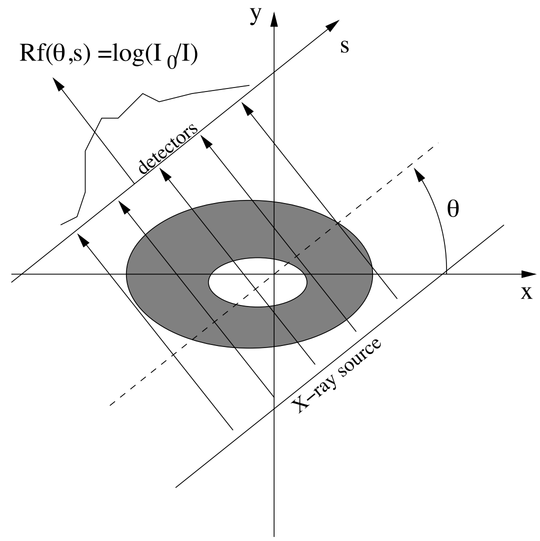

Figure 1: Principle of CT transmission tomography

transform can often be extended to these other transforms. This is one reason to study the Radon transform in this course. Another reason to study the Radon transform in a course on Fourier analysis is that the two transforms are closely related, as we will see. Two standard references are [6, 5]. These references were used to compile these notes, and some of the formulations of the theorems and proofs were taken from them.

2 Basic properties of the Radon transform

We first describe some properties of the Radon transform. Obviously $Rf(-\theta, -s) = Rf(\theta, s)$, i.e. Rf is an even function on the cylindrical subset

$$Z = S^{n-1} \times \mathbb{R} \tag{2}$$

of $\mathbb{R}^{n+1}$. We also write

$$R_\theta f(s) = Rf(\theta, s)$$

Denoting by $\theta^\perp$ the hyperplane normal to θ the formula for Rf can also be written as

$$\int_{\theta^\perp} f(s\theta + y)\, dy \tag{3}$$

If $f \in \mathscr{S}(\mathbb{R}^n)$, then Rf and $R_\theta f$ are in the Schwartz spaces on $\mathbb{R}^1$ and Z respectively, where the latter one is defined by using local coordinates or simply by restricting the function in $\mathscr{S}(\mathbb{R}^{n+1})$ to Z.

Some important properties of the Radon transform follow from formulas involving convolution and Fourier transforms. Whenever convolutions or Fourier transforms of functions on Z are used, they are to be taken with respect to the second variable, i.e. for $h, g \in \mathscr{S}(Z)$

$$h * g(\theta, s) = \int_{\mathbb{R}^1} h(\theta, s-t) g(\theta, t)\, dt,$$
$$\hat{h}(\theta, \sigma) = \int_{\mathbb{R}^1} e^{-is\sigma} h(\theta, s)\, ds.$$

The relationship between the Radon transform and the Fourier transform is of fundamental importance, and will later be used to prove the inversion formula. It is described in the following theorem, the so-called *Fourier slice theorem*

Theorem 1. *For $f \in \mathscr{S}(\mathbb{R}^n)$ we have*

$$\widehat{Rf}(\theta,\sigma) = \hat{f}(\sigma\theta), \qquad \sigma \in \mathbb{R}^1. \tag{4}$$

Proof. We have (with $(R_\theta f)\hat{}$ denoting the Fourier transform of $R_\theta f$)

$$\begin{aligned}(R_\theta f)\hat{}(\sigma) &= \int_{\mathbb{R}^1} e^{-i\sigma s} R_\theta f(s)\, ds \\ &= \int_{\mathbb{R}^1} e^{-i\sigma s} \int_{\theta^\perp} f(s\theta + y)\, dy\, ds.\end{aligned}$$

We now change integration variables. Instead of $(s,y) \in \mathbb{R} \times \theta^\perp$ we use $x = s\theta + y$, then $x \in \mathbb{R}^n$ and $s = \theta \cdot x$, $dx = dy\, ds$, hence

$$\begin{aligned}(R_\theta f)\hat{}(\sigma) &= \int_{\mathbb{R}^n} e^{-i\sigma\theta\cdot x} f(x)\, dx \\ &= \hat{f}(\sigma\theta).\end{aligned}$$

□

Next we will define a dual operator $R^\#$. We consider the inner product of Rf, $f \in \mathscr{S}(\mathbb{R}^n)$ and $g \in \mathscr{S}(Z)$. Using first the definition of the Radon transform, and then the change of variables $x = s\theta + y$, $s = \theta \cdot x$, $dx = dy\, ds$ we have

$$\begin{aligned}\int_{S^{n-1}} \int_{\mathbb{R}^1} Rf(\theta,s) g(\theta,s)\, ds\, d\theta &= \int_{S^{n-1}} \int_{\mathbb{R}^1} \int_{\theta^\perp} f(\theta s + y) g(\theta,s)\, dy\, ds\, d\theta \\ &= \int_{S^{n-1}} \int_{\mathbb{R}^n} f(x) g(\theta, \theta\cdot x)\, dx\, d\theta.\end{aligned}$$

Changing the order of integration we thus find that

$$\begin{aligned}\int_{S^{n-1}} \int_{\mathbb{R}^1} Rf(\theta,s) g(\theta,s)\, ds\, d\theta &= \int f(x) R^\# g(x)\, dx \\ R^\# g(x) &= \int_{S^{n-1}} g(\theta, x\cdot\theta)\, d\theta.\end{aligned}$$

Note that $R, R^\#$ also form a dual pair in the sense of integral geometry: while R integrates over all points in a plane, $R^\#$ integrates over all planes through a points.

Like for the Fourier transform we can determine the Radon transform of the derivatives of f in terms of the Radon transform of f, and similarly for the Radon transform of f multiplied by polynomials. Without proof we give the following results

$$R_\theta \partial_x^\alpha f = \theta^\alpha \partial_s^{|\alpha|} R_\theta f. \tag{5}$$

The Fourier slice theorem can also be used to prove the following convolution property.

$$R(f * g)(\theta,s) = \int_{\mathbb{R}^1} Rf(\theta,t) Rg(\theta, s-t)\, dt. \tag{6}$$

3 Inversion formulas

In this section we derive an explicit inversion formula for R. For $\alpha < n$ we define the linear operator I^α by

$$(I^\alpha f)\hat{}(\xi) = |\xi|^{-\alpha}\hat{f}(\xi).$$

I^α is called the Riesz potential. If I^α is applied to functions on Z it acts on the second variable. For $f \in \mathscr{S}$, $(I^\alpha f)\hat{} \in L_1(\mathbb{R}^n)$, hence $I^\alpha f$ makes sense and $I^{-\alpha}I^\alpha f = f$. For more on I^α, see [5], section V.5.

Theorem 2. *Let $f \in \mathscr{S}(\mathbb{R}^n)$ and let $g = Rf$. Then, for any $\alpha < n$, we have*

$$f = \tfrac{1}{2}(2\pi)^{1-n}\, I^{-\alpha} R^{\#} I^{\alpha-n+1} g. \tag{7}$$

Proof. We follow the proof from [6]. The Fourier inversion formula gives

$$I^\alpha f(x) = (2\pi)^{-n}\int_{\mathbb{R}^n} e^{ix\cdot\xi}|\xi|^{-\alpha}\hat{f}(\xi)\,d\xi.$$

Using polar coordinates $\xi = \sigma\theta$ this becomes

$$I^\alpha f(x) = (2\pi)^{-n}\int_{S^{n-1}}\int_0^\infty e^{i\sigma x\cdot\theta}|\sigma^{n-1-\alpha}|\hat{f}(\sigma\theta)\,d\sigma d\theta.$$

Using the Fourier slice theorem the Fourier transform $\hat{f}$ can be written in terms of $(Rf)\hat{}$, resulting in

$$I^\alpha f(x) = (2\pi)^{-n}\int_{S^{n-1}}\int_0^\infty e^{i\sigma x\cdot\theta}|\sigma|^{n-1-\alpha}(Rf)\hat{}(\theta,\sigma)\,d\sigma d\theta.$$

Replacing θ by $-\theta$ and σ by $-\sigma$ and using that $(Rf)\hat{}$ is even yields the same formula with the integral over $(0,\infty)$ replaced by the integral over $(-\infty,0)$. Adding both formulas leads to

$$I^\alpha f(x) = \tfrac{1}{2}(2\pi)^{-n}\int_{S^{n-1}}\int_{-\infty}^\infty e^{i\sigma x\cdot\theta}|\sigma|^{n-1-\alpha}(Rf)\hat{}(\theta,\sigma)\,d\sigma d\theta.$$

The inner integral can be expressed by the Riesz potential, hence

$$I^\alpha f(x) = \tfrac{1}{2}(2\pi)^{-n+1}\int_{S^{n-1}} I^{\alpha+1-n}Rf(\theta, x\cdot\theta)\,d\theta = \tfrac{1}{2}(2\pi)^{-n+1}R^{\#}I^{\alpha+1-n}Rf(x).$$

□

Equation (7) is a very general inversion formula. For all dimensions n it provides a family of inversion formulas parametrized by α. We want to make a couple of remarks, and look at some special cases.

1. Putting $\alpha = n-1$ in (7) yields

$$f = \frac{1}{2}(2\pi)^{1-n}I^{1-n}RR^{\#}g.$$

For n odd, I^{1-n} is simply a differential operator

$$I^{1-n} = (-\Delta)^{(n-1)/2}$$

In particular for $n = 3$ we obtain the formula

$$f(x) = -\frac{1}{8\pi^2}\Delta \int_{S^2} g(\theta, x \cdot \theta)\, d\theta, \tag{8}$$

where Δ acts on the variable x. This formula was already derived by Radon.

2. Putting $\alpha = 0$ in (7) we obtain

$$f = \frac{1}{2}(2\pi)^{1-n} R^{\#} I^{1-n} g, \tag{9}$$

where I^{1-n} acts on a function in $\mathbb{R}^1$. For $h \in \mathscr{S}(\mathbb{R}^1)$ we have

$$\begin{aligned}(I^{1-n}h)\hat{}(\sigma) &= |\sigma|^{n-1}\hat{h}(\sigma)\\ &= (\operatorname{sign}(\sigma))^{n-1}\sigma^{n-1}\hat{h}(\sigma).\end{aligned}$$

Multiplication by σ^{n-1} corresponds to apply $\left(-i\frac{d}{ds}\right)^{n-1}$. For odd n, the factor $(\operatorname{sign}(\sigma))^{n-1}$ is equal to one, and hence

$$f = \tfrac{1}{2}(2\pi)^{1-n}(-1)^{(n-1)/2} R^{\#} g^{(n-1)}(\theta, x\cdot\theta), \qquad n \text{ odd.}$$

For the factor $\operatorname{sign}(\sigma)$ that is present for even n, we note that the *Hilbert transform* can be defined by

$$(Hh)\hat{}(\sigma) = -i\operatorname{sign}(\sigma)\hat{h}(\sigma)$$

for $h \in \mathscr{S}(\mathbb{R}^1)$. The Hilbert transform is a well-known operator in Fourier analysis and signal processing. Because $\mathcal{F}\text{P.V.}\frac{1}{x} = -i\pi\operatorname{sign}(\xi)$ (see Grubb exercise 5.11), it can also be written as

$$\begin{aligned}(Hh)(s) &= \frac{1}{\pi}\left(\text{P. V. }\frac{1}{s}\right) * h(s)\\ &= \frac{1}{\pi}\text{ P. V.} \int_{-\infty}^{\infty} \frac{h(t)}{s-t}\, dt.\end{aligned} \tag{10}$$

We obtain

$$f = \tfrac{1}{2}(2\pi)^{1-n}(-1)^{(n-2)/2} R^{\#} H g^{(n-1)}(\theta, x\cdot\theta), \qquad n \text{ even.} \tag{11}$$

The Hilbert transform is non-trivial to compute numerically, due to the long tail of the function $\text{P.V.}\frac{1}{s}$, which has to be truncated becauses in practice only finite data is available. For numerical implementation the Fourier domain formula is a convenient choice. On a computer, a 1-D fast Fourier transform can be performed relatively easily (using the fast Fourier transform), and this leads to a very efficient implementation.

3. A filter is an operator that maps a function on $\mathbb{R}$ to a new function on $\mathbb{R}$ by convolution with a fixed function (the filter). The application of $R^{\#}$ is called backprojection. Formula (9) is therefore an example of *filtered backprojection*.

   In practice, formula (9) must be modified, because data is contaminated with noise. The operator I^{1-n} magnifies the large frequencies strongly (it is unbounded as an operator of L^2), in particular it magnifies the high-frequncy part of the noise, in such a way that the noise can become stronger than the correct signal. To address this, filtered backprojection formulas of the form

$$f \approx R^{\#}(w_b * g). \tag{12}$$

have been considered [6]. The function w_b is chosen to get a good approximate image by suppressing the large frequencies for which noise would be too large. For example it can contain a cutoff function $\Phi(|\sigma|/b)$, with b the cutoff frequency, i.e.

$$\hat{w}_b = \frac{1}{2}(2\pi)^{1-n}|\sigma|^{n-1}\Phi(|\sigma|/b).$$

Inversion formulas of the form (12) are called filtered backprojection.

4. In $n = 3$ dimensions, formula (8) shows that the inversion is *local*, i.e. the value of $f(x)$ only depends on the integrals of f over planes that intersect x, and a small neighborhood of these (due to the derivative operator). The same is true for all odd n. For n even this is not true, because the Radon transform present in (11) is not a local operator, as can be seen from (10).

4 The support theorem

We next consider the support theorem. Clearly, when f is supported in a ball with radius A, then $Rf(\theta, s) = 0$ for all $s > A$. The support theorem addresses the opposite implication. The theorem is as follows, for the proof see [5].

Theorem 3. *(The support theorem.) Let $f \in C(\mathbb{R}^n)$ satisfy the following conditions:*

(i) For each integer $k > 0$, $|x|^k f(x)$ is bounded.

(ii) There exists a constant $A > 0$ such that

$$Rf(\theta, s) = 0 \text{ for } |s| > A,$$

Then

$$f(x) = 0 \text{ for } |x| > A.$$

5 The range

It turns out that the functions in the range of the Radon transform satisfy certain relations, i.e. the range is not the entire space $\mathscr{S}(Z)$. The nature of these relations is described in the following lemma. (The material in this section is from Helgason [5], which contains a full treatment.)

Lemma 4. *For each $f \in \mathscr{S}(\mathbb{R}^n)$ the Radon transform $Rf(\theta, s)$ satisfies the following condition: For $k \in \mathbb{Z}_+$ the integral*

$$\int_{\mathbb{R}} Rf(\theta, s)s^k \, ds$$

can be written as a homogeneous polynomial of degree k in $\theta_1, \ldots, \theta_n$.

Proof. This is immediate from the relation

$$\begin{aligned}\int_{\mathbb{R}} Rf(\theta,s)s^k\,ds &= \int_{\mathbb{R}}\int_{\theta^\perp} s^k f(s\theta+y)\,dy\,ds\\ &= \int_{\mathbb{R}^n} f(x)(x\cdot\theta)^k\,dx. \qquad \square\end{aligned}$$

The conditions in Lemma 4 are called the Helgason-Ludwig consistency conditions. In accordance with this lemma we define the space

$$\mathscr{S}_H(Z) = \left\{ F\in\mathscr{S}(Z) \;\middle|\; \begin{array}{l}\text{For each } k\in\mathbb{Z}_+,\ \int_{\mathbb{R}} F(\theta,s)s^k\,ds \text{ is a homoge-}\\ \text{neous polynomial in } \theta_1,\dots,\theta_n \text{ of degree } k\end{array}\right\}.$$

We write

$$\mathscr{D}_H(Z) = C_0^\infty(Z)\cap\mathscr{S}_H(Z).$$

According to the lemma all functions in the range of R are part of $\mathscr{S}_H$, i.e.

$$R(\mathscr{S})\subset\mathscr{S}_H.$$

The next theorem states that actually equality holds.

Theorem 5. *The Radon transform $f\mapsto Rf$ is a linear one-to-one mapping of $\mathscr{S}(\mathbb{R}^n)$ onto $\mathscr{S}_H(Z)$.*

The proof can be found in [5].

6 Ill-posedness

In this section we give a brief introduction to ill-posedness, following chapter 4 of [6]. Let H, K be Hilbert space, and let A be a linear bounded operator in from H into K. We consider the problem

$$\text{given } g\in K,\text{ find } f\in H \text{ such that } Af=g, \tag{13}$$

i.e. the problem of the inversion of A.

The problem (13) is called well-posed by Hadamard [4] if it is uniquely solvable (i.e. A is bijective) for each $g\in K$ and if the solution depends continuously on g. Otherwise, (13) is called ill-posed. This means that for an ill-posed problem the operator A^{-1} does not exist, or is not defined on all of K, or is not continuous. The practical difficulty with an ill-posed problem is that even if its is solvable, the solution of $Af=g$ need not be close to the solution of $Af=g^\epsilon$ is g^ϵ is close to g.

In the sequel we will discuss how a meaningful solution to (13) can be defined. First we consider the case that A is not surjective and/or not injective using the so called Moore-Penrose generalized inverse. Then we will treat the case that the generalized inverse is not continuous (unbounded).

6.1 The Moore-Penrose generalized inverse

A pseudoinverse or generalized inverse is a linear map A^+, such that generalizes the inverse A^{-1}. When A is injective and surjective, the problem (13) can be solved for all g and we simply have $A^+ = A^{-1}$. When g is such that there is no f such that $Af = g$ we define A^+g as the function f that minimizes $\|Af - g\|$. This makes sense if $g \in \text{range}(A) + \text{range}(A)^\perp$. When there are multiple f that minimize $\|Af - g\|$, we take for A^+g the one with minimum norm. One can show that A^+ is a well-defined, possible unbounded, linear operator on $g \in \text{range}(A) + \text{range}(A)^\perp$.

Theorem 6. *$f = A^+g$ is the unique solution of*

$$A^*Af = A^*g \tag{14}$$

in $\overline{\text{range}(A^*)}$.

Proof. f minimizes $\|Af - g\|$ if and only $(Af - g, Au) = 0$ for all $u \in H$, i.e. if and only if $A^*Af = A^*g$. Among all solutions of this equation the unique element with least norm is the one in $(\ker(A))^\perp = \overline{\text{range}(A^*)}$. □

We remark that the system $A^*Af = A^*g$ is usually called the normal equations. When A is an $m \times n$ matrix, the operator A^*A is invertible if and only if A is injective. Hence we must have $m \geq n$ and the columns must be linearly independent.

6.2 The singular value decomposition

We recall that the singular value decomposition of an $m \times n$ matrix M is a factorization

$$M = U\Sigma V^*,$$

where U is an $m \times m$ unitary matrix, V is an $n \times n$ unitary matrix, and Σ is an $m \times n$ diagonal matrix with real, non-negative entries on the diagonal. It is convention to order the entries of Σ in decreasing order.

When we denote the columns of U by u_k, the columns of V by v_k and the values on the diagonal of Σ by σ_k, this implies that Mx can be written as

$$Mx = \sum_{k=1}^{\min(m,n)} \sigma_k(x, v_k)u_k.$$

Using a slightly different notation, this will be our definition for the singular value decomposition in the case of L^2 spaces, i.e. by a singular value decomposition we mean a representation of A in the form

$$Af = \sum_{k=1}^{\infty} \sigma_k(f, f_k)g_k, \tag{15}$$

where (f_k), (g_k) are normalized orthogonal systems in H, K respectively and σ_k are positive numbers, the singular values of A. We always assume the sequence $\{\sigma_k\}$ to be bounded. Then, A is a linear continuous operator from H into K with adjoint

$$A^*g = \sum_{k=1}^{\infty} \sigma_k(g, g_k)f_k, \tag{16}$$

and the operators

$$A^*Af = \sum_{k=1}^{\infty} \sigma_k^2 (f, f_k) f_k, \tag{17}$$

$$AA^*g = \sum_{k=1}^{\infty} \sigma_k^2 (g, g_k) g_k \tag{18}$$

are self-adjoint operators in H, K respectively. The spectrum of A^*A consists of the eigenvalues of σ_k^2 with eigenelements f_k and possibly of the eigenvalue 0 whose multiplicity may be infinite. The same is true for AA^* with eigenelements g_k. The two eigensystems are related by

$$A^* g_k = \sigma_k f_k, \qquad A f_k = \sigma_k g_k. \tag{19}$$

Vice versa, if (f_k), (g_k) are normalized eigensystems of A^*A, AA^* respectively such that (19) holds, then A has the singular value decomposition (15). In particular, compact operatorsalways admit a singular value decomposition.

Theorem 7. *If A has the singular value decomposition (15), then*

$$A^+ g = \sum_{k=1}^{\infty} \sigma_k^{-1} (g, g_k) f_k. \tag{20}$$

For the finite-dimensional case is easily proved. For the infinite-dimensional we refer to [6].

6.3 Ill-posedness and regularization

The presence of zero singular values in a matrix or operator A implies that this operator is not injective and/or not surjective. However, also the presence of small singular values leads to problems. From (23) it follows that if $\sigma_k \to 0$, then A^+ is unbounded, hence not continuous.

Let g^ϵ be an approximation to g such that $\|g - g^\epsilon\| \leq \epsilon$. Knowing only g^ϵ, we can say that $|(g, g_k) - (g^\epsilon, g_k)| < \epsilon$. If formula (23) is used with g^ϵ instead of g, the error in the k-th coefficient satisfies

$$|\sigma_k^{-1}(g, g_k) - \sigma_k^{-1}(g^\epsilon, g_k)| \leq \epsilon / \sigma_k. \tag{21}$$

Since the right hand side becomes large for σ_k small, the contribution of g_k to A^+ cannot be computed reliably for such k. Thus looking at the singular values and the corresponding elements g_k shows which features of the solution f of (13) can be determined from an approximation g^ϵ to g and which can not!

This problem is handled by modifying the operator A^+, depending on ϵ, using so called *regularization.* A regularization of A^+ is a family of linear continuous operators $T_\gamma : K \to H$ which are defined on all of K and for which

$$\lim_{\gamma \to 0} T_\gamma g = A^+ g \tag{22}$$

on the domain of A^+. Regularization is a large topic, of which we just discuss a small part.

Obviously, $\|T_\gamma\| \to \infty$ as $\gamma \to 0$ if A^+ is not bounded. With the help of a regularization we can solve (13) approximately in the following sense. Let $g^\epsilon \in K$ be an approximation to g such that $\|g^\epsilon - g\| \leq \epsilon$. Let $\gamma(\epsilon)$ be such that, as $\epsilon \to 0$,

$$\text{(i) } \gamma(\epsilon) \to 0 \qquad \text{(ii) } \|T_{\gamma(\epsilon)}\|\epsilon \to 0.$$

Then, as $\epsilon \to 0$,

$$\begin{aligned} \|T_{\gamma(\epsilon)}g^\epsilon - A^+g\| &\leq \|T_{\gamma(\epsilon)}(g^\epsilon - g)\| + \|T_{\gamma(\epsilon)}g - A^+g\| \\ &\leq \|T_{\gamma(\epsilon)}\|\epsilon + \|T_{\gamma(\epsilon)}g - A^+g\| \\ &\to 0 \end{aligned}$$

Hence, $T_{\gamma(\epsilon)}g^\epsilon$ is close to A^+g if g^ϵ is close to g.

The number γ is called a regularization parameter. Determining a good regularization parameter is one of the crucial points in the application of regularization methods. We will not discuss this matter. We rather assume that we can find a good regularization parameter by trial and error.

There are several methods for constructing a regularization.

6.4 The truncated singular value decomposition

We saw above that the coefficients k for which σ_k is small are poorly determined if g is replaced by g^ϵ. A regularization method is to simply omit these coefficients. The truncated singular value decomposition T_γ is defined by

$$A^+g = \sum_{k \leq 1/\gamma} \sigma_k^{-1}(g, g_k)f_k. \tag{23}$$

If follows from theorem 7 that $T_\gamma g \to A^+g$ as $\gamma \to 0$, and T_γ is bounded with $\|T_\gamma\| \leq \sup_{k \leq q/\gamma} \sigma_k^{-1}$.

6.5 The method of Tikhonov-Phillips

For completeness we mention the method of Tikhonov-Phillips, often simply called the Tikhonov method. Here we put

$$T_\gamma = (A^*A + \gamma I)^{-1}A^* = A^*(A^*A + \gamma I)^{-1}. \tag{24}$$

Equivalently, $f_\gamma = T_\gamma g$ can be defined to be the minimizer of

$$\|Af - g\|^2 + \gamma\|f\|^2. \tag{25}$$

The operator $(A^*A+\gamma I)$ symmetric positive definite and its spectrum is contained in $[\gamma, C]$ for some real constant C. By functional analysis it follows that is has a bounded inverse. Using further arguments from functional analysis it also follows that $T_\gamma g \to A^+g$ is $g \in D(A^+)$.

7 Singular value decomposition of the Radon transform

In this section we study the singular value decomposition of the Radon transform. It was established by Davison [2], for the general case of functions of n variables, assuming that both the domain and the image space are suitable weighted spaces of square-integrable functions. Here we consider the case of $n = 2$ dimensions, and we consider the most natural choice for the weight in the image space, following [1].

We will first show the continuity of R between these weighted L_2 spaces. Then we can determine the singular value decomposition. This will involve a set of special functions (see equation (30) below) that forms a basis for a certain L_2 space. Such special functions perform a similar role as the functions $h_n(x) = e^{inx}$, $\ldots, -2, -1, 0, 1, 2, \ldots$ in Fourier analysis (the h_n form a basis for $L_2([0, 2\pi])$ with periodic boundary conditions).

From the singular value decomposition we will see that the inversion problem for the Radon transform is indeed ill-posed, because we will see that $\sigma_k \to 0$. (Constructing regularized inversion operators falls outside the scope of these notes.)

7.1 Continuity of the Radon transform on L_2 spaces

Let Ω be the unit disk in $\mathbb{R}^2$. We will consider the Radon transform on functions f supported in Ω. This means that Rf supported in $Z_1 \stackrel{\text{def}}{=} S^1 \times [-1, 1]$. Let

$$w(s) = (1 - s^2)^{1/2},$$

which is such that the length of the intersection of a line $x \cdot \theta = s$ with the unit disk is given by $2w(s)$. By $L^2(Z_1, w(s)^{-1})$ we denote the weighted L_2 space with norm

$$\|g\|^2_{L_2(Z,q(\theta,s))} = \int_{S^1} \int_{-1}^{1} |g(\theta, s)^2| w(s)^{-1}\, ds\, d\theta.$$

It turns out that this is quite a natural norm for the range of R. We have the following

Theorem 8. *Let $n = 2$. The operator*

$$R : L_2(\Omega) \to L_2(Z_1, (1 - s^2)^{-1/2})$$

is continuous and

$$\|Rf\|_{L_2(Z_1,(1-s^2)^{-1/2})} \leq \sqrt{4\pi} \|f\|_{L_2(\Omega)}. \tag{26}$$

Proof. To start we will assume f is continuous and establish the estimate (26). It then follows that R can be extended to $f \in L_2(\Omega)$.

The Radon transform can be written as

$$(Rf)(\theta, s) = \int_{-w(s)}^{w(s)} f(s\theta + t\theta^\perp)\, dt, \qquad |s| \leq 1. \tag{27}$$

The right hand side can be viewed as the L^2 inner product of the function $1_{[-w(s),w(s)]}$ and the function f. Applying the Cauchy-Schwarz inequality we find

$$|(Rf)(\theta, s)|^2 \leq 2w(s) \int_{-w(s)}^{w(s)} |f(s\theta + t\theta^\perp)|^2\, dt.$$

so that

$$\begin{aligned}\int_{-1}^{1} w(s)^{-1}|(Rf)(\theta,s)|^2\,ds &\le 2\int_{-1}^{1}\int_{-w(s)}^{w(s)} |f(s\theta+t\theta^\perp)|^2\,dt\,ds \\ &= 2\int_\Omega |f(x)|^2\,dx.\end{aligned} \tag{28}$$

Integrating over θ we find (26). □

7.2 The singular value decomposition

To determine the SVD we look for the eigenvalues of RR^*, making use of (17). Like is common in the study of eigenvalues of partial differential operators, we use separation of variables, i.e. we assume $g(\theta,s)$ is of the form

$$g(\theta,s) = S(s)\Theta(\theta), \tag{29}$$

and try to find eigenfunctions of this form.

When performing integrations over $\theta\in S^1$ we will typically parametrize this as $\theta = (\cos(\phi),\sin(\phi))$ and integrate over ϕ from 0 to 2π.

Let $U_m(s)$ be the Chebyshev polynomials of the second kind, which are defined by

$$U_m(s) = \frac{\sin[(m+1)\arccos s]}{\sin(\arccos s)}.$$

These polynomials are orthogonal with respect to the weight function $w(s)$; more precisely they satisfy the following orthogonality and normalization conditions

$$\int_{-1}^{1} w(s)U_m(s)U_{m'}(s)\,ds = \frac{\pi}{2}\delta_{m,m'}.$$

Furthermore, the $U_m(s)$ form an orthogonal basis of the weighted L_2 space $L_2([-1,1],w(s))$. It follows from this that

$$\text{the function } U_m(s)w(s)^{-1} \text{ form a basis for } L_2([-1,1],w(s)^{-1}). \tag{30}$$

Let $\mathcal{Z} = L_2(Z_1,(1-s^2)^{-1/2})$. We consider the subspaces $\mathcal{Z}_m$ which are defined as the spaces containing the functions

$$g_m(\theta,s) = \sqrt{\frac{2}{\pi}}w(s)U_m(s)u(\theta), \qquad m=0,1,\ldots \tag{31}$$

where $u(\theta)\in L^2(S^1)$. It turns out that the spaces $\mathcal{Z}_m$ are invariant under RR^*, If g_m is a function in $\mathcal{Z}_m$, then, from the formulas for R and R^* we obtain

$$(RR^*g_m)(\theta,s) = \sqrt{\frac{2}{\pi}}\int_{-w(s)}^{w(s)}\int_0^{2\pi} U_m\big[\theta'\cdot(s\theta+t\theta^\perp)\big]u(\theta')\,d\phi'\,dt \tag{32}$$

which we will show leads to

$$(RR^*g_m)(\theta,s) = \frac{4\pi}{m+1}\sqrt{\frac{2}{\pi}}w(s)U_m(s)\bar{u}(\theta), \tag{33}$$

where

$$\bar{u}(\theta) = \frac{1}{2\pi}\int_0^{2\pi} \frac{\sin[(m+1)(\phi-\phi')]}{\sin(\phi-\phi')} u(\theta')\, d\phi'. \tag{34}$$

To prove (33) and (34) we exchange the order of integration in (32) and we consider the integral

$$I(s,\theta,\theta') = \int_{-w(s)}^{w(s)} U_m\left[\theta' \cdot (s\theta + t\theta^{\perp})\right] dt.$$

If $\theta = (\cos\phi, \sin\phi)$ and $\theta' = (\cos\phi', \sin\phi')$, we have $\theta{\cdot}\theta' = \cos(\phi-\phi')$ and $\theta^{\perp}{\cdot}\theta' = -\sin(\phi-\phi')$. Then, if we write $s = \cos\xi$ and $\psi = \phi - \phi'$, we obtain

$$I(\cos\xi,\theta,\theta') = \int_{\sin\xi}^{\sin\xi} U_m(\cos\xi\cos\psi - t\sin\psi)\, dt.$$

By the change of variable $u = \cos\xi\cos\psi - t\sin\psi$ this integral becomes

$$I(\cos\xi,\theta,\theta') = \int_{\cos(\xi+\psi)}^{\cos(\xi-\psi)} \frac{U_m(u)}{\sin(\psi)}\, du$$

and by introducing the new variable $u = \cos\eta$ we finally obtain

$$\begin{aligned} I(\cos\xi,\theta,\theta') &= \frac{1}{\sin\psi}\int_{\xi-\psi}^{\xi+\psi} \sin[(m+1)\eta]\, d\eta \\ &= \frac{2}{m+1}\frac{\sin[(m+1)\psi]}{\sin\psi}\sin[(m+1)\xi]. \end{aligned}$$

By substituting this expression in equation (32), with $\psi = \phi - \phi'$ and $\xi = \arccos s$, we get equations (33) and (34).

Next we must determine the eigenfunctions of the map $u \mapsto \bar{u}$ given in (34). A basis for $L^2(S^1)$ is given by the functions

$$Y_l(\theta) = \frac{1}{\sqrt{2\pi}} e^{-il\phi}$$

It is straightforward to check that if $u = Y_l$ then

$$\bar{u} = \begin{cases} Y_l & \text{if } -m \le l \le m \text{ and } l-m \text{ is even} \\ 0 & \text{otherwise.} \end{cases}$$

Summarizing we obtain the following. Because the $U_m(s)w(s)^{-1}$ form a basis for $L^2([-1,1], w(s)^{-1})$ and the Y_l form a basis for $L_2(S^1)$, the functions

$$w(s)^{-1}U_m(s)Y_l(\theta) \text{ form a basis of } L_2(Z_1, w(s)^{-1}),$$

with $m = 0, 1, \ldots$, and $l \in \mathbb{Z}$. We define

$$u_{m,k}(\theta,s) = w(s)^{-1}U_m(s)Y_{m-2k}(\theta)$$

then these are the eigenfunctions of RR^* with nonzero eigenvalues

$$RR^* u_{m,k} = \sigma_m^2 u_{m,k}, \qquad k = 0, 1, \ldots, m$$

where

$$\sigma_m = \left(\frac{4\pi}{m+1}\right)^{1/2}.$$

The functions

$$v_{m,k} = \frac{1}{\sigma_m} R^* u_{m,k}$$

are the other part of the SVD. Without proof we note that they are given by

$$v_{m,k}(x) = (2m+2)^{1/2} Q_{m,|m-2k|}(|x|) Y_{m-2k}\left(\frac{x}{|x|}\right).$$

where

$$Q_{m,l}(r) = r^l P^{(0,l)}_{\frac{1}{2}(m+l)}(2r^2 - 1)$$

$P_n^{(\alpha,\beta)}(t)$ being the Jacobi polynomial of degree n.

We see that $\sigma_m \to 0$, so that the inverse problem for the Radon transform is indeed ill-posed. We also observe that RR^* has many zero eigen values, corresponding to the fact that R is not surjective (cf. Theorem 5).

The general case of n dimensions involves more knowledge of special functions (e.g. Gegenbauer polynomials) and their properties, and can be found in [2, 6].

References

[1] M. Bertero and P. Boccacci. *Introduction to inverse problems in imaging.* Institute of Physics Publishing, Bristol, 1998.

[2] M. E. Davison. A singular value decomposition for the Radon transform in n-dimensional Euclidean space. *Numerical Functional Analysis and Optimization*, 3(3):321–340, 1981.

[3] C. L. Epstein. *Introduction to the mathematics of medical imaging.* Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, second edition, 2008.

[4] J. Hadamard. Le problème de Cauchy et les équations aux dérivées partielles linéaires hyperboliques. Paris: Hermann & Cie. 542 pp. Frs.100.00 (1932)., 1932.

[5] S. Helgason. *The Radon transform*, volume 5 of *Progress in Mathematics.* Birkhäuser Boston, Inc., Boston, MA, second edition, 1999.

[6] F. Natterer. *The mathematics of computerized tomography*, volume 32 of *Classics in Applied Mathematics.* Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2001. Reprint of the 1986 original.