Deconvolution and Regularization for Numerical Solutions of Incorrectly Posed Problems

M. Iqbal
Deconvolution and Regularization for Numerical Solutions of Incorrectly Posed Problems

M. Iqbal
Department of Mathematical Sciences
King Fahd University of Petroleum and Minerals
Dhahran 31261, Saudi Arabia
e-mail: miqbal@kfupm.edu.sa

Abstract

In this paper we have converted the Laplace transform into Fredholm integral equation of the first kind of convolution type, which is an ill-posed problem and used deconvolution and regularization method to solve it. The method is applied to various test examples taken from the literature.

The method gives a good approximation to the true solution. Our results confirm that our method is competitive with other recently presented methods. The results are shown in the table and the diagrams.

AMS (MOS) subject classification: 65R20, 65R30

Key words: Ill-posed problems, Inversion of Laplace transform, convolution equation, deconvolution, regularization, filter function, regularization parameter.
1 Introduction

The Laplace transform inversion is a severely ill-posed problem in the terminology of improperly posed problems. Unfortunately, many problems of physical interest lead to Laplace transforms whose inverses are not readily expressed in terms of tabulated functions, although there exist extensive tables of transforms and their inverses. It is highly desirable, therefore, to have methods for approximate numerical inversion. However, no single method gives optimum results for all purposes and all occasions. For a detailed bibliography, the reader should consult Piessens [24], Piessens and Branders [25], and a review and comparison is given in Davies [6] and Talbot [28].

The problem of the recovery of a real function $f(t)$ for $t \geq 0$, given its Laplace transform

$$\int_0^\infty e^{-st} f(t) dt = g(s) \quad (1.1)$$

for real values of $s$, is an ill-posed problem in the sense of Hadamard and is therefore affected by numerical instability. The ill-posedness of Laplace transform inversion in the case where $f \in L^2(R_+)$ and $g(s)$ is known for all real and positive values of $s$, can be investigated by means of the Mellin transform [18, 29].
The term incorrectly posed or improperly posed means that the solution $f(t)$ of equation (1.1), may not be unique, may not exist and may not depend continuously on the data.

Incorrectly posed inverse problems have become a recurrent theme in modern sciences, for example, crystallography [11], geophysics [1], medical electrocardiograms [10], metrology [27], radio astronomy [14], reservoir engineering [15] and tomography [32]. Corresponding to this broad spectrum of fields of applications, there is a wide literature on different kinds of inversion algorithms for evaluating the inverse problems.

✓ The basic principle, common to all such methods is as follows: Seek a solution that is consistent both with observed data and prior notions about the physical behavior of the phenomenon under study. Different authors have employed different methods such as the method of regularization [31,34], maximum entropy [20], quasi reversibility [15], and cross validation [18,34]. Regularization methods have been discussed by Varah [31], Essa and Delves [9], Wahba [33,34] Eggermont [7], Thompson [30], Ang [2], Rudolf [26], Bertero [3], and Brianzi [4].

In practice, however $g(s)$ is known only in a finite set of points. The case of an infinite set of equidistant points was investigated by Papoulis [23]. Regularization methods have been discussed by Varah [31], Essah and Delves [9] and Brianzi [4]. Several other methods have been developed [7,8,12,15,16,19,22,35] which, in general,
work rather well even if they require a large computational cost and high precision arithmetic.

In (1.1) given \( g(s) \) for \( s \geq 0 \) we wish to find \( f(t) \), \( t \geq 0 \) and \( f(t) = 0 \) for \( t < 0 \), so that (1.1) holds. Frequently \( g(s) \) is measured at certain points. We assume \( g(s) \) is given analytically with known \( f(t) \) so that we can measure the error in the numerical solution.

2 Fredholm Equation of Convolution Type

We shall convert Laplace transform into the first kind integral equation of convolution type, with the following substitution in equation (1.1)

\[
s = a^x \text{ and } t = a^{-y} \text{ where } a > 1
\]  

(2.1)

Then

\[
g(a^x) = \int_{-\infty}^{\infty} \log a e^{-a^x - y} f(a^{-y}) a^{-y} dy
\]  

(2.2)

Multiplying both sides of (2.2) by \( a^x \) we obtain the convolution equation

\[
\int_{-\infty}^{\infty} K(x - y) F(y) dy = G(x), \quad -\infty \leq x \leq \infty
\]  

(2.3)

\[
\begin{align*}
G(x) &= a^x g(a^x) = sg(s) \\
K(x) &= \log a a^x e^{-a^x} = \log a se^{-s} \\
F(y) &= f(a^{-y}) = f(t)
\end{align*}
\]  

(2.4)
Equation (2.3) occurs widely in the applied sciences. $K$ and $G$ are known kernel and data functions respectively, and $F$ is to be found. We shall assume that $F, G$ and $K$ lie in suitable function spaces, such as $L_2(R)$, so that their Fourier transform ($FT_\ast$) exist. ($\wedge$ denotes $FT_\ast$ and $\vee$ denotes inverse $FT_\ast$).

3 Description of the Method

We assume that the support of each function $F, G$ and $K$ is essentially finite and contained within the interval $[0, T]$, where $T$ is the period and $T = Nh, N$ is the number of data points and $h$ is the spacing. Let $T_N$ denote the space of trigonometric polynomials of degree at most $N$ and period $T$. We shall look for filtered solution of (2.3) within the space $T_N$ for the following reasons:

(a) The descretization error in the convolution may be made precisely zero at the grid points.

(b) Fast Fourier Transform (FFT) routines are easily employed in the solution procedure.

(c) The adoption of $T_N$ as the approximating function space is itself a regularizing feature.

Let $G$ and $K$ be given at $N$ equally spaced points $x_n = nh, n = 0, 1, 2, \ldots N - 1$, with spacing $h = T/N$. Then $G$ and $K$ are interpolated by $G_N$ and $K_N \in T_N$.
where

\[ G_N(x) = \frac{1}{N} \sum_{q=0}^{N-1} \hat{G}_{N,q} \exp(i\omega_q x) \quad (3.1) \]

\[ \hat{G}_{N,q} = \sum_{n=0}^{N-1} G_n \exp(-i\omega_q x_n) \quad (3.2) \]

where

\[ G(x_n) = G_n = G_N(x_n), \quad \omega_q = \frac{2\pi q}{T} \quad (3.3) \]

Similar expressions as (3.1) and (3.2) can be obtained for \( K_N \).

Consider (2.3) the Fredholm integral equation of the first kind of convolution type

\[(KF)(x) \equiv \int_{-\infty}^{\infty} K(x - y)F(y)dy = G(x), \quad -\infty < y < \infty, \quad (3.4)\]

where \( G \) and \( K \) are known functions in \( L_2(R) \), and \( F \in H^p(R) \) is to be found. Then from the convolution we have

\[ \hat{K}(\omega) \hat{F}(\omega) = \hat{G}(\omega) \quad (3.5) \]

where

\[ F(y) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{\hat{G}(\omega)}{\hat{K}(\omega)} \exp(i\omega y)d\omega. \quad (3.6) \]

The ill-posedness of (3.4) is reflected by the fact that any small perturbation \( \epsilon \) in \( G \), whose transform \( \hat{\epsilon}(\omega) \) does not decay faster than \( \hat{K}(\omega) \) as \( |\omega| \to \infty \) will result in a perturbation in \( \frac{\hat{G}(\omega)}{\hat{K}(\omega)} \) which will grow without bound.
The ill-posed nature of the integral equation can be easily understood if we take a Fourier transform of both sides of equation (3.4) then we have equation (3.5) in that equation \( \hat{F}(\omega), \hat{K}(\omega) \) and \( \hat{G}(\omega) \) are the Fourier transforms of \( F(y), K(x) \) and \( G(x) \) respectively.

It can be shown that the Fourier 'Image' of the Laplace Kernel \( K(x-y) = \exp(-st) \) is band limited i.e. \( \hat{K}(\omega) \) decreases to zero as \( \frac{1}{\omega^2 + \xi^2} \) for \( \omega \to \infty \). The Laplace operator can thus be compared with a low pass filter in electronics. Using this analogy, one can say that the high frequency components of the Fourier spectrum \( \hat{g}(\omega) \) of \( g(s) \) are cut off by the band limited Laplace integral operator if \( \omega > \omega_{\text{max}} \), where \( \omega_{\text{max}} \) is a certain threshold frequency.

When \( G \) is inexact, therefore, we may seek a stable or filtered approximation to \( F \) given by

\[
F_\lambda(y) = \frac{1}{2\pi} \int_{-\infty}^{\infty} Z(\omega; \lambda) \frac{\hat{G}(\omega)}{\hat{K}(\omega)} \exp(i\omega y) d\omega
\]

(3.7)

where \( Z(\omega; \lambda) \) is a stabilizing or filter function dependent on a parameter \( \lambda \).

In this paper we restrict attention to filters generated from regularization theory.

The smoothing functional

\[
C(F; \lambda) = ||KF - G||^2_2 + \lambda \Omega[F]
\]

(3.8)
is minimized in an appropriate subspace of $L_2$, where $\Omega[F]$ is a stabilizing functional in the form of a smoothing norm

$$\Omega[F] = ||LF||^2$$  \hspace{1cm} (3.9)

and $L$ is a linear operator. The regularization parameter $\lambda$ controls the trade-off between smoothness, as imposed by $\Omega$ and the extent to which (3.4) is satisfied.

In this paper we construct a method which determines $\lambda$ optimally. In the case of numerical deconvolution, at least the method compares extremely well with other methods available in the literature.

We restrict attention to regularization of order $p$ ($p = 2$ in our case), where $L$ in (3.9) is the $p$-th order differential operator, $LF = F^{(p)}$ and the norm in (3.9) is $L_2$. The minimizer of (3.8) in $H^p$ is then given by (3.7) where

$$z(\omega; \lambda) = \frac{|\hat{K}(\omega)|^2}{|\hat{K}(\omega)|^2 + \lambda \omega^{2p}}$$  \hspace{1cm} (3.10)

The equation (3.4) is now replaced by

$$(K_N F_N)(x) \equiv \int_0^T K_N(x - y)F_N(y)dy = G_N(x)$$  \hspace{1cm} (3.11)

where $K_N$ is periodically continued outside $(0, T)$. Then we may prove (a) above.
Lemma 3.1. Let \( F \in T_N \) and \( F = (F(x_0),\ldots,F(x_{N-1}))^T \in R^N \). Then the \( N \times N \) matrix

\[
K = \psi \text{diag} \left( \hat{K}_{N,q} \right) \psi^H
\]

(3.12)

where \( \psi \) is the unitary matrix with elements

\[
\psi_{rs} = \frac{1}{\sqrt{N}} \exp \left( \frac{2\pi i rs}{N} \right), \quad r, s = 0,\ldots,N - 1
\]

(3.13)

has the property

\[
(KF)_n = (K_N F)(x_n).
\]

(3.14)

Thus from the infinite support hypothesis and (3.3) it follows that at \( \{ x_n \} \), (2.3) is exactly equivalent to the discrete system

\[
(KF)_n = G_n
\]

(3.15)

where \( K \) is given in (3.12) and \( F = (F_N(x_0),\ldots,F_N(x_{N-1}))^T \).

In \( T_N \) it is easily shown that \( F_\lambda \) in (3.7) is approximated by

\[
F_{N,\lambda}(x) = \sum_{q=0}^{N-1} Z_{q,\lambda} \frac{\hat{G}_{n,q}}{\hat{K}_{N,q}} \exp(i\omega_q x)
\]

(3.16)

where the discrete \( p \)-th order filter is

\[
Z_{q,\lambda} = \frac{\left| \hat{K}_{N,q} \right|^2}{\left| \hat{K}_{N,q} \right|^2 + N^2 \lambda \omega_q^{2p}},
\]

(3.17)
and

\[ \tilde{\omega}_q = \begin{cases} \omega_q, & 0 \leq q < \frac{1}{2} N \\ \omega_{N-q}, & \frac{1}{2} N \leq q < N - 1 \end{cases} \]  

(3.18)

To show (b) above we note that $\sqrt{N}\psi^{HT}$ is the discrete FT matrix representing (3.2), and so (3.15) is equivalent to the diagonal system

\[ \hat{K}_{N,q} \hat{F}_{n,q} = \hat{G}_{N,q} \]  

(3.19)

After regularization, (3.19) is replaced by

\[ \hat{K}_{N,q} \hat{F}_{n,q;\lambda} = Z_{q;\lambda} \hat{G}_{N,q} \]  

(3.20)

so that $F_{N;q;\lambda}(x)$ may be found by multiplying the FF of $\{G_N\}$ by the filter, dividing by the FFT of $\{K_n\}$, and then taking the inverse FFT (fast Fourier transform).

4 Choice of Regularization Parameter

Suppose we ignore the $j$-th data point $G_j$, and define filtered solution $F_{N;\lambda}^{(j)}(x) \in T_N(T_N$ stands for trigonometric polynomial of degree $N$) as the minimizer of

\[ \sum_{n=0}^{N-1} [(K_N \ast F) (x_n) - G (x_n)]^2 + \lambda ||F^{(P)}(x)||^2. \]

Then we get a vector $G_{N;\lambda}^{(j)} \in R_N$ defined by

\[ G_{N;\lambda}^{(j)} = K F_{N;\lambda}^{(j)} \]  

(4.1)
Now the $j$-th element $G_{N,\lambda,j}^{(j)}$ of equation (4.1) should predict the missing value $G_j$. We may thus construct the weighted mean square prediction error over all $j$, i.e.,

$$V(\lambda) = \frac{1}{N} \sum_{j=0}^{N-1} Q_j(\lambda) \left[ G_{N,\lambda,j}^{(j)} - G_j \right]^2$$

(4.2)

The filtered solution to the problem should minimize the mean square prediction error in (4.2). To minimize $V(\lambda)$ in the form given by equation (4.2) is a time-consuming process.

Wahba [32] has suggested an alternative expression which depends on a particular choice of weights and results in considerable simplification.

Let

$$E_{N,\lambda} = (F_{N,\lambda}(x_0), F_{N,\lambda}(x_1), \ldots, F_{N,\lambda}(x_{N-1}))^T$$

(4.3)

and define

$$G_{N,\lambda} = KE_{N,\lambda}.$$

Then there exists a matrix $A(\lambda)$, called an influence matrix such that

$$G_{N,\lambda} = A(\lambda)G_N$$

(4.4)

Let $K = \text{diag} \left( \hat{K}_{N,q} \right)$ and $\hat{Z} = \text{diag} \left( Z_{q,\lambda} \right)$. Then from (4.1), we have

$$E_{N,\lambda} = \psi(\hat{K})^{-1} \hat{Z} \hat{G}_N$$

where

11
\[ \hat{G}_N = \psi^H G_N \text{ and so} \]
\[ A(\lambda) = \psi \hat{Z} \psi^H. \] (4.5)

Also
\[ K = \psi \hat{K} \psi^H \]

where \( \psi \) is the unitary matrix with elements
\[ \psi_{rs} = \frac{1}{\sqrt{N}} \exp \left[ \frac{2\pi i}{N} rs \right], \ r, s = 0, 1, \ldots, N - 1. \] (4.6)

Wahba has shown in a more general context that the weights must be
\[ Q_j(\lambda) = \left[ \frac{1 - a_{jj}(\lambda)}{\frac{1}{N} \text{Trace}(1 - A(\lambda))} \right]^2, \ j = 0, 1, \ldots, N - 1 \] (4.7)

where \( A(\lambda) \) is the influence matrix in (4.4). An estimate for \( \lambda \) is found by Wahba [32].

This enables the expression (4.2) to be written as
\[ V(\lambda) = \frac{\frac{1}{N} \| (1 - A(\lambda)) \hat{G}_N \|^2}{\left[ \frac{1}{N} \text{Trace}(1 - A(\lambda)) \right]^2}. \] (4.8)

Using (4.5), we get
\[ V(\lambda) = \frac{\frac{1}{N} \| (1 - \hat{\varepsilon}) \hat{G}_N \|^2}{\left[ \frac{1}{N} \text{Trace}(1 - \hat{\varepsilon}) \right]^2} \]
i.e.,
\[ V(\lambda) = \frac{\frac{1}{N} \sum_{q=0}^{N-1} (1 - z_q \lambda)^2 |\hat{G}_{n,q}|^2}{\left[ \frac{1}{N} \sum_{q=0}^{n-1} (1 - z_q \lambda) \right]^2}. \] (4.9)
The expression in (4.9) is minimized using the quadratic interpolation technique to obtain a minimum. Natterer [19] derives an optimal convergence rate for $\lambda = \lambda(\delta)$ for a range of values of $n$. It was Wahba [32], who first observed that for discrete problems, the critical factor affecting convergence rates is $\sigma^2/n$, where $\sigma^2$ is the variance.

5 Deconvolution and Optimal Filtering

The term optimal is to be found in several different contexts in the literature on regularization. The term optimal filter is often used synonymously with the minimum variance or Wiener filter, the latter being reserved usually for a stochastic setting (the stochastic regularization of the deconvolution problem).

Under some mild conditions of $K$, we then have a definite singular system

$$
\begin{align*}
K_n V_{ni} &= k_{ni} u_{ni} \\
K_n^* u_{ni} &= k_{ni} v_{ni} 
\end{align*}
$$

$i = 1, 2, \ldots, n$ \hspace{1cm} (5.1)

with singular vectors $u_{ni} \in \mathbb{R}^n$, singular function $v_{ni} \in H$.

In the discrete setting of (3.7) with orthogonal singular vectors normalized by

$$
\langle u_{ni}, u_{nj} \rangle = \frac{1}{n} (u_{ni}^T u_{nj}) = \delta_{ij}
$$

a minimum variance filter defined by

$$
Z^*_i, i = 1, 2, \ldots n, \hspace{1cm} (5.2)
$$
where
\[ Z_i^* = Z(k, \lambda) = \frac{k^2}{(k^2 + \lambda)} \] (5.3)
which minimizes
\[ E \left( \frac{1}{n} \left| \sum_{i=1}^{n} Z_i (y_i, U_{ni}) U_{ni} - G \right|^2 \right), \] (5.4)
(where \( y_i = G(x_i) + \delta_i \quad i = 1, 2, 3, \ldots n \)) over all sequences \( Z_i \).

\( E \) denotes expectation with respect to the noise distribution. It can be shown that
\[ Z_i^* = \frac{(G, u_{ni})^2}{(G, u_{ni})^2 + \sigma^2/n} \] (5.5)

Alternatively, if the (continuous) data function \( G \) and noise \( \delta \) are modelled by independent stationary stochastic processes with spectral densities \( P_G \) and \( P_\delta \), respectively, then the stochastic equivalent of (5.4) is minimized by the Wiener filter
\[ Z^*(i) = \frac{P_G(i)}{P_G(i) + P_\delta(i)} \] (5.6)
where \( i \), here, denotes a spectral variable which may be continuous.

Natterer [21] concludes that “there is nothing wrong with higher order regularization (we have \( p = 2 \) in this paper), even well above the order of smoothness of the exact solution, the only mistake one can make is to regularize with an order which is too low”.
6 Numerical Examples

In this section we tabulate the results of the above method applied to test problems available in the literature.

All data functions have the property $g(s) = 0(s^{-1})$ and since it is a severally ill-posed problem, therefore, no noise is added apart from machine rounding error.

In all cases we have taken $N = 256$ data points.

**Example 1.** This example has been taken from Theocaris ([29], case 5, page 79)

$$f(t) = e^{-at} \quad \alpha = 1.0$$

$$g(s) = \frac{1}{s + \alpha}$$

The numerical calculations are given in Table 1, and diag (1). The optimal solution is compared with Theocaris solution in diag (4).

**Example 2.** This example has been taken from Theocaris ([29, case 4, page 79].

$$f(t) = e^{-at} \sin(\beta t)$$

where $\alpha = 5.0$, $\beta = 2.2$

$$g(s) = \frac{u \sin v}{s^2 + 2us \cos v + u^2}, \quad \left\{ \begin{array}{l}
u = (\alpha^2 + \beta^2)^{-\frac{1}{2}} \\
v = \tan^{-1}(\beta/\alpha)
\end{array} \right.$$
The numerical calculations are given in table 1 and diag (2). The optimal solution is compared with Theocaris solution in diag (5).

**Example 3.** This example has been taken from Brianzi and Mcwhiter ([4, 19]),

\[ f(t) = t^\alpha e^{-\beta t} \quad \text{for} \quad \alpha = 1.0, \quad \beta = 1.0 \]

\[ g(s) = \frac{\Gamma(\alpha)}{(s + \beta)^{\alpha+1}} \]

The numerical calculations are given in Table 1 and diag (3). The optimal solution is compared with Mcwhirter's solution in diag (6).

7 Conclusion

Our method worked well over all the three test examples. The results obtained are shown in diags (1-3). The Theocaris and Mcwhirter's solutions are also presented in diags (4-6), for comparison purposes. Our results are better and are shown over a wider range of \( t \), than the results of Theocaries and Mcwhirter.

8 Acknowledgement

The author acknowledges the excellent research and computer facilities availed at King Fahd University of Petroleum and Minerals, Dhahran, during the preparation of this paper.
9 Choice of $T$ and $h$

In order to solve (3.3) we need to choose two numbers $x_{min}$ and $x_{max}$ such that $|G(x)| < \varepsilon$ whenever $x < x_{min}$ and $x > x_{max}$. In what follows we choose $\varepsilon = (10)^{-4}(\max |G(x)|)$. We find $x_{min}$ and $x_{max}$ as the smallest and largest solutions of the non-linear equation $G(x) = \varepsilon$, we may then pose the deconvolution problem (3.3) on the interval $[0, T]$, where $T = x_{max} - x_{min}$. Since the size of the essential support of $G(x)$ depends upon 'a', we write $T = T_a$ for a fixed number $N$ of equidistant data points $\{x_n\}$, we have spacing $h = \frac{T}{N}$.

<table>
<thead>
<tr>
<th>Example</th>
<th>T</th>
<th>h</th>
<th>a</th>
<th>$\lambda$</th>
<th>$V(\lambda)$</th>
<th>$|f - f_\lambda|_\infty$</th>
<th>Diag</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>12.50</td>
<td>0.04883</td>
<td>10.0</td>
<td>$0.9601 \times 10^{-9}$</td>
<td>$0.7323 \times 10^3$</td>
<td>0.005</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>9.0</td>
<td>0.03516</td>
<td>5.0</td>
<td>$0.109 \times 10^{-10}$</td>
<td>$0.81315 \times 10^4$</td>
<td>0.0992</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>12.01</td>
<td>0.04727</td>
<td>10.0</td>
<td>$0.192 \times 10^{-10}$</td>
<td>$0.1042 \times 10^2$</td>
<td>0.001</td>
<td>3</td>
</tr>
</tbody>
</table>
References


Diag (1) Solution by Regularization Method

\[ f(t) \]

\[ t \]

- Num.Soln for \( \alpha = 10.0 \)
- True Soln.

Diag (2) Solution by Regularization Method

\[ f(t) \]

\[ t \]

- Num.Soln for \( \alpha = 5.0 \)
- True Soln.

Diag (3) Solution by Regularization Method

\[ f(t) \]

\[ t \]

- Num.Soln for \( \alpha = 10.0 \)
- True Soln.
Figure 5. $p_w(n)$ as a function of $v$, $\beta = 4\pi$ and noise $-10^{-7}$; broken curve, $N = 20$; full curve, actual solution $p(v) = v e^{-v}$ (and $N = 40$ and $N = 60$).