

# New approaches to the optimal regularization

Takashi Kitagawa

Institute of Information Sciences and Electronics  
University of Tsukuba

March 8, 1992

## Abstract

This paper introduces two new approaches to determine the optimal parameter in the method of regularization. One is based on the error analysis made in [4] and [5]. The other is based on, what is called in [2], L-curve, which is formulated and analyzed in [3].

## 1 Introduction

One of the most important problems in approximating the solution of a linear ill-posed problems by the method of regularization resides in the selection of the optimal regularization parameter. we present new two approaches to the optimal regularization.

We consider the ill-conditioned linear systems arising from Fredholm integral equations of the first kind of the form

$$\int_a^b k(s, t) \hat{f}(t) dt = \hat{g}(s), \quad s_{min} \leq s \leq s_{max}, \quad (1)$$

where  $K(s, t)$  and  $\hat{g}(s)$  are known  $L_2$  functions and  $\hat{f}$  is the unknown function in  $L_2[a, b]$ . This equation is known to be an ill-posed problem in the sense that  $\hat{f}$  does not depend on  $\hat{g}$  continuously, namely, any small perturbation in  $\hat{g}$  results in arbitrarily large change in  $\hat{f}$ . Via some discretization process, one can reduce (1) to the equation

$$T f = g, \quad (2)$$

with  $f = (f_1, f_2, \dots, f_n) \in R^n$ ,  $g = (g_1, g_2, \dots, g_m) \in R^m$  and  $T : R^n \mapsto R^m$ .

The ill-posedness of (1) results from the fact that the operator  $\hat{T}$  which is the integral operator in (1) does not have a bounded inverse, which in turn, implies that the condition number of the matrix  $T$  increases rapidly as  $m$  and  $n$  increase. Consequently, any attempts to solve (2) by a conventional least squares method may produce disastrous results. A number of methods are available to mitigate the effect of this ill-conditioning. Best known of them are the truncation of the singular value decomposition and the method of regularization.

## 2 Optimal regularization

The method of regularization solves the related well-posed problem of minimizing a smoothing functional. In other words:

For given  $g_\Delta = g + \Delta g \in R^m$ , find  $f = f(\mu, \Delta g) \in R^n$  and  $\mu \in [0, \infty)$  for which

$$\min_{f \in R^n} \{ \|Tf - g_\Delta\|^2 + \mu \|f\|^2 \} \quad (3)$$

is attained.

The parameter  $\mu$  is called the regularization parameter, which controls the tradeoff between the stability of the system (3) and the fidelity to the original equation. This technique is known to be very successful in practice, provided that the optimal value of the regularization parameter  $\mu$  is determined appropriately [1, 4, 6].

We set, for further use,

$$e(\mu; \Delta g) = T^\dagger g - f(\mu; \Delta g), \quad (4)$$

where  $T^\dagger$  denotes the Moore-Penrose generalized inverse of  $T$  and  $f(\mu; \Delta g)$  represents the minimizer of the smoothing functional (3).

We define the optimal regularization parameter as follows.

**Definition 1** We call  $\mu_o$  the optimal regularization parameter if

$$\mu_o \in \{ \bar{\mu} \mid \min_{\mu \in [0, \infty)} \|e(\mu)\| = \|e(\bar{\mu})\| \}. \quad (5)$$

Hereafter we may write  $f(\mu) = f(\mu; \Delta g)$ , etc. for simplicity.

### 3 New approaches to the optimal regularization

We present the following two new approaches to this problem:

1) The first approach is by introducing a function to determine the optimal parameter. The method chooses the value of  $\mu$  for which

$$\min_{\mu \in P_\sigma} \zeta(\mu) \quad \text{with} \quad \zeta(\mu) = \left\| \frac{d}{d\xi} f(\mu; \Delta g) \right\| \quad (6)$$

is attained, where  $P_\sigma$  is the set of singular values of  $T^t T$  and  $\xi = \log \mu$ . We monitor the values of the function  $\zeta(\mu)$  among the values of  $\sigma_i^2$ 's, where  $\sigma_i$ ,  $i = 1, 2, \dots, n$ , are singular values of  $T$ . Then we employ the value of  $\mu$  which gives the minimum of  $\zeta(\mu)$ . Namely, one advantage of this method is that the number of the evaluation of the function is at most  $n$ . The theoretical aspect which explains why this method works out well is discussed in [4] and the practical numerical algorithm together with some numerical experiments are given in [5].

2) The second approach uses the notion of L-curve which is termed by [2] and is defined as the graph of

$$(\|r_\mu^\Delta\|, \|f(\mu)\|) \quad \text{with} \quad r_\mu^\Delta = T f(\mu) - g_\Delta \quad (7)$$

which is parametrized by  $\mu$ . The name of L-curve comes from the numerical observation that the graph (7) has a steep bend in its middle and it looks like L. Moreover, the corner of the L-curve gives a good estimation for the optimal regularization parameter  $\mu_0$ .

The maximizer of the curvature of the L-curve is employed as the optimal parameter. The formulation of this method with numerical examples is given in [3]. The explicit expression of  $\kappa(\mu)$  using the singular system, which is not simple at all but we can compute anyway, is given as follows;

$$\kappa(\mu) = \frac{1}{(\|r_\mu^\Delta\|^2 + \mu^2 \|f(\mu)\|^2)^{\frac{3}{2}}} \left| \frac{\|r_\mu^\Delta\|^2 \|f(\mu)\|^2 (\Sigma_1(\mu) + 3\mu \Sigma_2(\mu))}{\Sigma_3(\mu)^2} - \mu (\mu \|r_\mu^\Delta\|^2 + \|f(\mu)\|^2) \right| \quad (8)$$

where

$$\Sigma_1(\mu) \equiv \sum_{i=1}^k \frac{\sigma_i^2(\sigma_i^2 - 2\mu)}{(\sigma_i^2 + \mu)^4} (u_i, g_\Delta)_m^2 \quad (9)$$

$$\Sigma_2(\mu) \equiv \sum_{i=1}^k \frac{\sigma_i^2}{(\sigma_i^2 + \mu)^4} (u_i, g_\Delta)_m^2 \quad (10)$$

$$\Sigma_3(\mu) \equiv \sum_{i=1}^k \frac{\sigma_i^2}{(\sigma_i^2 + \mu)^3} (u_i, g_\Delta)_m^2 \quad (11)$$

with  $u_i$ 's are the left singular vectors of  $T$  and  $k = \text{rank}(T)$ .

## References

- [1] Groetsche, C. W., *The Theory of Tikonov Regularization for Fredholm Integral Equation of the First Kind*, Pitman, Boston, 1984.
- [2] Hansen, P. C., *Analysis of discrete ill-posed problems by means of the L-curve*, preprint 1990.
- [3] Hosoda, Y. and Kitagawa, T., *Optimal regularization for ill-posed problems by means of the L-curve*, Trans. of JSIAM, **2**(1992), (in Japanese).
- [4] Kitagawa, T., *A deterministic approach to the optimal regularization - the finite dimensional case -*, Japan J. of Appl. Math. **4**(1987), pp.371-379
- [5] Kitagawa, T., *A numerical method to estimate the optimal regularization parameter*, J. of Info. Proc., **11**(1988), pp.263-270
- [6] Nashed, M. Z., *Operator theoretic and computational approaches to ill-posed problems with applications to antenna theory*, IEEE Trans. on Antennas and Propagat., **29**(1981), pp.220-231.