

# ON OPTIMAL REGULARIZATION

V. Ya. Arsenin, V. V. Ivanov

1968

SovietRxiv

---

View the original and related papers at <https://sovietrxiv.org/items/ru-196801.51804>

Source: Math-Net.Ru and CyberLeninka. Machine translation. Verify with the original.

## Abstract

## Full Text

UDC 517.948.32/33

MATHEMATICS

V. Ya. ARSENIN, V. V. IVANOV

# ON OPTIMAL REGULARIZATION

(Presented by Academician A. N. Tikhonov, 3 I 1968)

1. We consider integral equations of the first kind of convolution type

$$Kz \equiv \int_{-\infty}^t k(t-t')z(t') dt' = u(t), \quad (1)$$

the problem of solving which is ill-posed. Usually the right-hand side  $u(t)$  is known only approximately with some accuracy  $\delta$  (in  $L_2$ ),  $u(t) = u_T(t) + v(t)$ , where  $u_T(t)$  is the exact value,  $v(t)$  is an error (noise) and is a random quantity. We assume that the solution  $z(t)$ , the right-hand side, and the noise are realizations of stationary random processes and that  $z(t)$  and  $v(t)$  are uncorrelated. By  $N(\omega)$  and  $S(\omega)$  we shall denote the spectral densities of the solution  $z(t)$  and of the noise  $v(t)$ .

The works <sup>(2,3)</sup> contain the following formulation of the problem, analogous to the problem of optimal Wiener filtering <sup>(1)</sup>. Let  $\{Q\}$  be some family of operators defined on all functions  $u \in L_2(-\infty, \infty)$ , and let  $B$  be a given operator with domain of definition  $U_1 \subset L_2(-\infty, \infty)$ . Let  $u_T(t) \in U_1$ .

**Problem 1.** Find in the class  $\{Q\}$  such an operator  $Q_0$  that

$$\overline{(Q_0 u - Bu_T)^2} = \min_{\{Q\}} \overline{(Qu - Bu_T)^2}. \quad (2)$$

The operator  $Q_0$  is called the optimal operator in  $\{Q\}$  in the sense of (1). In application to solving equation (1),  $B \equiv K^{-1}$ . If in this case the family of operators  $\{Q\}$  consists of linear integral operators of convolution type, then Problem 1 is a problem of optimal filtering. In this case the operator  $Q_0$  is determined by the spectral densities  $N(\omega)$  and  $S(\omega)$  <sup>(1-3)</sup>.

If we do not have information about the spectral densities  $N(\omega)$  and  $S(\omega)$ , then to find an approximate solution of equation (1) (but no longer an optimal one in the sense of (2)) one may use the regularization method <sup>(4)</sup>. In this case a family of regularizing linear operators depending on the parameter  $\alpha$ ,  $Q_{\alpha, M}$ , is constructed, determined with the aid of the regularizer

$$\Omega[z] = \int_{-\infty}^{\infty} M(\omega) |z(\omega)|^2 d\omega,$$

where  $z(\omega)$  is the Fourier transform of the function  $z(t)$  from the class of possible solutions;  $M(\omega)$  is a given even nonnegative function such that  $M(0) \geq 0$ ,  $M(\omega) > 0$  for  $\omega \neq 0$ , and, for sufficiently large  $|\omega|$ ,  $M(\omega) \geq c > 0$ . The regularized approximate solutions  $z_\alpha(t)$  are determined by the formula

$$z_\alpha(t) = Q_{\alpha,M} u = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \frac{k(-\omega)u(\omega)}{L(\omega) + \alpha M(\omega)} e^{i\omega t} d\omega, \quad L(\omega) = k(-\omega)k(\omega), \quad (3)$$

where  $k(\omega)$  and  $u(\omega)$  are the Fourier transforms of the kernel  $k(t)$  and of the right-hand side  $u(t)$  of equation (1). Formula (3) defines a one-parameter family of linear regularizing operators  $Q_{\alpha,M}$ , on which one may seek the solution of problem I with  $B \equiv K^{-1}$ . This family is determined by specifying the function  $M(\omega)$ .

**Theorem 1.** *There exists a one-parameter family of regularizing operators  $Q_{\alpha,M}$  on which the solution of problem I coincides with the solution obtained by the method of optimal filtering.*

Such a family is determined by the function  $M(\omega)$ , equal, up to a numerical factor, to the ratio  $S(\omega)/N(\omega)$ . The parameter  $\alpha$  is found from the condition  $\min_{\alpha} \overline{(Q_{\alpha,M} u - z)^2}$  ( $z(t)$  will be called the signal).

2. In computational practice one usually uses regularizers of the Tikhonov type of order  $p$  (5), i.e., one specifies the function  $M(\omega)$  in the form  $M(\omega) = \omega^{2p}$ , which determines a family of regularizing operators  $Q_{\alpha,p}$ , and, using some algorithm, determines the parameter  $\alpha$  (6).

We shall 1) show that there exists a solution of problem I in the class  $Q_{\alpha,p}$  of Tikhonov regularizing operators; 2) indicate the relation between the optimal value of the regularization parameter  $\alpha$  and the high-frequency characteristics of the signal and noise under the following assumptions: a) the Fourier transform of the kernel  $k(t)$ , i.e.  $k(\omega)$ , on the real axis for large  $|\omega|$  has an asymptotic form  $S(\omega) = S_0/\omega^{2r}$ ,  $N(\omega) = N_0/\omega^{2n}$ , where  $S_0, N_0$  are constants,  $S_0 \geq 0$ ,  $m > 1$ ) (see (7)); b) the spectral densities of the functions  $z(t)$  and  $v(t)$ , i.e.  $N(\omega)$  and  $S(\omega)$ , on the real axis for large  $|\omega|$  have asymptotic form  $S(\omega) = S_0/\omega^{2r}$ ,  $N(\omega) = N_0/\omega^{2n}$ , where  $S_0, N_0$  are constants,  $S_0 \geq 0$ ,  $N_0 > 0$ . For  $r = 0$  we have "white" noise, for  $r = 1$  "thermodynamic" noise. The number  $n$  coincides with the order of smoothness of the function  $z(t)$ . We assume that  $S_0$  is small and  $0 < \gamma, \mu < 1$ , where  $\gamma = (2n - 1)/(2p + 2k)$ ,  $\mu = (2k - 2r + 1)/(2p + 2k)$ . For Tikhonov regularizing operators problem I is formulated as follows:

**Problem I'.** Among the regularizing operators  $Q_{\alpha,p}$  of Tikhonov type, find an operator  $Q_{\alpha_0,p_0}$  such that

$$\overline{(Q_{\alpha_0 p_0} u - z)^2} = \min_{\alpha, p} \overline{(Q_{\alpha, p} u - z)^2}.$$

Here  $z(t)$  is the solution of equation (1) with right-hand side  $u_T(t)$ . The operator  $Q_{\alpha_0 p_0}$  will be called the optimal regularizing operator of Tikhonov type (of order  $p_0$ ), and the function  $z_{\alpha_0}(t) = Q_{\alpha_0 p_0} u$  the optimal, in the sense of  $I'$ , regularized solution of equation (1). Let  $\xi = Q_{\alpha, p} v$  and  $T(\alpha, p) = \overline{(Q_{\alpha, p} u - z)^2}$ .

**Theorem 2.** For kernels I (under the condition  $\mu > 0$ ) and II types, the function  $T(\alpha, p)$  as a function of  $\alpha$  attains its smallest value at  $\alpha = \alpha_{r.o.}$ , determined by the formulas:

$$\alpha_{r.o.} = A^2 \left\{ \frac{S_0}{A^2 N_0} \frac{\mu(1-\mu) \sin \gamma \pi}{\gamma(1-\gamma) \sin \mu \pi} \right\}^{1/(\gamma+\mu)} \quad \text{for kernels of type I;} \quad (4)$$

$$\alpha_{r.o.} = \frac{S_0}{N_0} \omega_1^{2n-2p-2r} \quad \text{for kernels of type II.} \quad (5)$$

Here  $\omega_1$  is the root of the equation

$$\omega(-2L + \alpha\omega^{2p}) dL/d\omega - 2(p+n)L\alpha\omega^{2p} + 2(2p-n)L^2 = 0. \quad (6)$$

$$T(\alpha, p) = \overline{\xi^2} + \overline{(\delta z_\alpha)^2}, \quad (7)$$

where  $\delta z_\alpha = Q_{\alpha, p} u_T - z(t)$ . It is easy to find that

$$\overline{\xi^2} = \frac{1}{\pi} \int_0^\infty \frac{S(\omega)L(\omega) d\omega}{(L(\omega) + \alpha\omega^{2p})^2}, \quad \overline{(\delta z_\alpha)^2} = \frac{1}{\pi} \int_0^\infty \frac{\alpha^2 \omega^{4p} N(\omega) d\omega}{(L(\omega) + \alpha\omega^{2p})^2}. \quad (8)$$

$\alpha_{r.o.}$  is found from the condition  $\partial T/\partial \alpha = 0$ .

Using the smallness of  $S_0$  (and  $\alpha$ ), the integrands may be replaced by their asymptotic values. Obviously,  $\alpha_{p.o.} = \alpha(p)$ . Therefore  $T(\alpha_{p.o.}, p) = \psi(p)$ .

3. The quality of the approximate solution  $z_{\alpha_{p.o.}}(t)$ , obtained under optimal regularization of order  $p$  of Tikhonov type, can be characterized by the quantity  $\psi(p)$ . Let us consider the dependence of  $\psi(p)$  on the regularization order  $p$ . Denote by  $z_{o.f.}(t)$  the solution of equation (1) by the optimal filtering method. This is the best, in the sense of (2), approximate solution of equation (1). For equations (1) with a kernel of type I or II,

$$\overline{[z_{o.f.}(t) - z(t)]^2} = \psi(p_0) + O(S_0).$$

**Theorem 3.** For kernels of type I the function  $\psi(p)$  has the following properties:

- 1) it has a minimum at  $p = p_0 = n - r$ ;
- 2) it increases monotonically for  $p > p_0$  and tends to a finite limit  $\psi(\infty)$  as  $p \rightarrow \infty$ ;

$$3) \quad \frac{\psi(\infty)}{\psi(p_0)} = \frac{x^2}{\pi(x-1)} \sin\left(\frac{\pi}{x}\right) + \frac{1}{x}O(S_0) < \frac{4}{\pi} + \frac{1}{x}O(S_0); \quad (9)$$

4)

$$\frac{\psi''(p_0)}{\psi(p_0)} = \frac{y(1-y)}{(2n-1)^2} \left\{ \frac{(\pi y)^2}{\sin^2(\pi y)} - 1 - \frac{y^2}{1-y^2} \right\} + yO(S_0), \quad y = \frac{r}{\mu},$$

$$x = 1 + \frac{1}{y}. \quad (10)$$

The proof of these properties is obtained by estimating the integrals (8) using formula (4).

**Theorem 4.** For kernels of type II the function  $\psi(p)$  has the following properties: 1) it has a minimum at  $p = p_0 = n - r$ ; 2)  $\psi(p) = mN_0/\omega_1^{2n-1}(A_1\omega_1)^{1/m}$ ; 3)  $\psi''(p_0)/\psi(p_0) = 6(A_1\omega_1)^{-2/m}$ . Here  $\omega_1$  is the root of equation (6),  $A_1 = A_0(\cos \pi/2m)^m$ .

The proof is carried out analogously to the case of a kernel of type I.

**Corollary.** For strongly smoothing operators  $K$  (an operator with a kernel of type II and an operator with a kernel of type I, if  $k \gg n, r$ ), the mean square deviation of the regularized solution of equation (1) from the exact solution depends only weakly on the order  $p$  of the regularizer.

Indeed, a measure of the sensitivity of the function  $\psi(p)$  to the order  $p$  of the regularizer near  $p = p_0$  is the ratio  $\psi''(p_0)/\psi(p_0)$ . As is seen from (10), for  $k \gg n, r$ , i.e. for  $y$  close to zero, this ratio is small.

4. Let us consider the question of how much the optimal regularized solutions  $z_{\alpha_m, p_m}(t)$ , obtained with the aid of regularizing operators of different orders  $p_m$ , differ from one another. A measure of the deviation is the quantity

$$R(p_1, p_2) = \overline{(z_{\alpha_1, p_1}(t) - z_{\alpha_2, p_2}(t))^2}.$$

Obviously,

$$R(p_1, p_2) \leq \psi(p_1) + \psi(p_2) + O(S_0).$$

For small  $k$  this estimate is sufficiently good. For strongly smoothing operators this estimate can be refined.

**Theorem 5.** For strongly smoothing operators  $K$  with kernels of type I or II,

$$R(p_1, p_2) = O[\psi(p_1) - \psi(p_2)] + O[\psi(p_2) - \psi(p_0)],$$

where on the right the order of smallness is taken with respect to the variable  $1/k$ .

**Proof.** We find directly that

$$R(p, p_0) = \frac{1}{\pi} \int_0^\infty \frac{LN(\alpha\omega^{2p} - S/N)^2 d\omega}{(L + \alpha\omega^{2p})^2(L + S/N)}.$$

The functions  $(\alpha\omega^{2p}/(L + \alpha\omega^{2p}))^2$  and  $L/(L + S/N)$  may be replaced respectively by the unit functions  $\eta(\omega - \tilde{\omega}_1)$  and  $\eta(\omega_2 - \omega)$ , where  $\tilde{\omega}_1$  is the root of the equation  $L(\omega) = \alpha\omega^{2p}$ , and  $\omega_2$  is the root of  $L(\omega) = S(\omega)/N(\omega)$ . In this case  $\omega_2 > \tilde{\omega}_1$ , if  $p > p_0$ . (It suffices to consider only this case.) Therefore, in the formula for  $R(p, p_0)$  it is necessary to integrate only over the interval  $(\tilde{\omega}_1, \omega_2)$ . Esti-

this integral, we obtain  $R(p, p_0) = O[\psi(p) - \psi(p_0)]$ . From this estimate and from the inequality  $R(p_1, p_2) \leq R(p_1, p_0) + R(p_2, p_0)$  the theorem follows.

Thus, for the indicated classes of equations, in finding an approximate optimal solution of equation (1), instead of optimal filtering one may use the regularization method with regularizers of Tikhonov type, whose order has only a weak effect on the result.

5. To compute  $a_{r.o.}$ , as is seen from formulas (4)–(5), one must know the high-frequency characteristics of the signal and the noise  $N_0, n, S_0, r$ . Below we shall show that, for ergodic processes, the indicated parameters can be uniquely determined from the family of regularized solutions  $\{z_\alpha(t)\}$ , corresponding to various values of  $\alpha$  and obtained with the aid of regularizing operators of Tikhonov type  $Q_{\alpha,p}$ , whose order  $p$  can be chosen sufficiently arbitrarily. This means that, when applied to ergodic processes, the regularization method makes it possible to find the optimal approximate solution of equation (1) (or one close to it), using substantially less information about the sought solution and the noise than in the method of optimal filtering.

**Theorem 6.** *The function*

$$F(\alpha) = \alpha^{-2k/(k+p)} \overline{(\alpha z_\alpha^{(2p)})^2}$$

for small noises (i.e., for small  $S_0$ ) has a minimum at some point  $\alpha_{\min}$  and a maximum at some point  $\alpha_{\max}$ . Moreover,  $\alpha_{\min}/\alpha_{\max} \rightarrow 0$  as  $S_0 \rightarrow 0$ .

**Proof.**

$$F(\alpha) = \frac{1}{\pi} \alpha^{2p/(k+p)} \int_0^\infty \omega^{4p} L(LN + S)/(L + \alpha\omega^{2p})^2 d\omega.$$

For kernels of type I, for small  $\alpha$

$$F(\alpha) = O(\alpha^{2p/(k+p)}) + c_1\alpha^\nu + c_2\alpha^{-\mu},$$

where

$$c_1 = \frac{\pi N_0}{4q^2} \frac{2p - 2k - 2n + 1}{\sin(4k + 2n - 1)\pi/2q}, \quad c_2 = \frac{\pi S_0}{A^2 4q^2} \frac{2p - 2r + 1}{\sin(2p - 2r + 1)\pi/2q},$$

$q = k + p$ . Consequently, as  $\alpha \rightarrow 0$ ,  $F(\alpha) \rightarrow \infty^*$ . For large  $\alpha$ ,

$$F(\alpha) = O(\alpha^{-2k/q}).$$

Hence the theorem follows for kernels of type I. For kernels of type II the proof is carried out analogously.

**Remark 1.** The condition  $\alpha_{\min}/\alpha_{\max} \ll 1$  may serve as a criterion of smallness of the noise and of the applicability of the results presented to the analysis of experimental curves  $u(t)$ .

**Remark 2.** In the case of ergodic processes, the function  $F(\alpha)$  is determined without a priori knowledge of the high-frequency characteristics of the sought solution  $z(t)$  and the noise, by the formula

$$F(\alpha) = \alpha^{-2k/q} \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T \{K^*(Kz_\alpha - u)\}^2 dt,$$

where  $K^*$  is the operator adjoint to the operator  $K$ .

**Theorem 7.** *The high-frequency characteristics of the sought solution and of the noise  $N_0, n, S_0, r$  are uniquely determined by the behavior of the function  $F(\alpha)$ .*

**Proof.** On the interval  $(0, \alpha_{\min})$  the curve  $y = F(\alpha)$  is determined only by the parameters  $r, S_0, p, k$ , and on the interval  $(\alpha_{\min}, \alpha_{\max})$  by the parameters  $N_0, n, p, k$ . These parameters can be determined from the indicated portions of the curve  $y = F(\alpha)$  with arbitrary accuracy.

We express our deep gratitude to A. N. Tikhonov for interesting discussions.

Received 29 XII 1967

## CITED LITERATURE

1. N. Wiener, *Extrapolation, Interpolation and Smoothing of Stationary Time Series*, N. Y., 1949.
2. L. A. Vainshtein, V. D. Zubakov, *Extraction of Signals Against a Background of Random Interference*, Moscow, 1960.

3. M. M. Lavrent'ev, V. G. Vasil'ev, *Siberian Mathematical Journal*, 7, No. 3 (1966).
4. A. N. Tikhonov, *Dokl. Akad. Nauk SSSR*, 151, No. 3, 501 (1963).
5. A. N. Tikhonov, *Dokl. Akad. Nauk SSSR*, 153, No. 1, 49 (1963).
6. A. N. Tikhonov, V. B. Glasko, *Zhurnal Vychislitel'noi Matematiki i Matematicheskoi Fiziki*, 4, No. 2, 569 (1964).
7. V. Ya. Arsenin, V. V. Ivanov, *Zhurnal Vychislitel'noi Matematiki i Matematicheskoi Fiziki*, 8, No. 2 (1968).

\* In a neighborhood of  $\alpha_{\min}$ , for small  $S_0$  and large  $p$  ( $4p > 2n - 1$ ),  $O(\alpha^{2p/q})$  is much smaller than the term  $c_1\alpha^\nu$ .

*Note: Figure translations are in progress. See original paper for figures.*

*Source: Math-Net.Ru and CyberLeninka. Machine translation. Verify with the original.*