Continuous Regularization Methods

Aditya Ganeshan

Indian Institute of Technology, Roorkee

December 13, 2016

Table of Content

A-Priori Parameter Choice Rules

We will consider the class of linear regularization methods based on spectral theory for self-adjoint operators.

Let $\{E_\lambda\}$ be the spectral family of T^*T. If T^*T is continuously invertible, then the best approximate solution, $x^\dagger = T^\dagger y$ can be written as

$$x^\dagger = \int \frac{1}{\lambda} dE_\lambda T^* y \tag{1}$$

A-Priori Parameter Choice Rules

We will consider the class of linear regularization methods based on spectral theory for self-adjoint operators.

Let $\{E_\lambda\}$ be the spectral family of T^*T. If T^*T is continuously invertible, then the best approximate solution, $x^\dagger = T^\dagger y$ can be written as

$$x^\dagger = \int \frac{1}{\lambda} dE_\lambda T^* y \tag{1}$$

Now, if $Tx = y$ is ill-posed, then the integral does not exist, since the integrand $1/\lambda$ has a pole in 0, which belongs to the spectrum of T^*T.

Now, for regularization we replace the integrand $1/\lambda$ by a parameter-dependent family of functions $g_\alpha(\lambda)$ which are

- at least piece wise continuous on $[0, \|T\|^2]$.
- continuous for the right in points of discontinuity.

Now, if $Tx = y$ is ill-posed, then the integral does not exist, since the integrand $1/\lambda$ has a pole in 0, which belongs to the spectrum of T^*T.

Now, for regularization we replace the integrand $1/\lambda$ by a parameter-dependent family of functions $g_\alpha(\lambda)$ which are

- at least piece wise continuous on $[0, \|T\|^2]$.
- continuous for the right in points of discontinuity.

This gives us,

$$x_\alpha = \int g_\alpha(\lambda) dE_\lambda T^* y, \tag{2}$$

and, for noisy data y^δ with $\|y^\delta - y\| < \delta$

$$x_\alpha^\delta = \int g_\alpha(\lambda) dE_\lambda T^* y^\delta. \tag{3}$$

Therefore we can write the regularization methods $\{R_\alpha\}$ as,

$$R_\alpha = \int g_\alpha(\lambda) dE_\lambda T^*. \tag{4}$$

This gives us,

$$x_\alpha = \int g_\alpha(\lambda) dE_\lambda T^* y, \tag{2}$$

and, for noisy data y^δ with $\|y^\delta - y\| < \delta$

$$x_\alpha^\delta = \int g_\alpha(\lambda) dE_\lambda T^* y^\delta. \tag{3}$$

Therefore we can write the regularization methods $\{R_\alpha\}$ as,

$$R_\alpha = \int g_\alpha(\lambda) dE_\lambda T^*. \tag{4}$$

This gives us,

$$x_\alpha = \int g_\alpha(\lambda) dE_\lambda T^* y, \tag{2}$$

and, for noisy data y^δ with $\|y^\delta - y\| < \delta$

$$x_\alpha^\delta = \int g_\alpha(\lambda) dE_\lambda T^* y^\delta. \tag{3}$$

Therefore we can write the regularization methods $\{R_\alpha\}$ as,

$$R_\alpha = \int g_\alpha(\lambda) dE_\lambda T^*. \tag{4}$$

Now for the residual $x^\dagger - x_\alpha$ we have,

$$x^\dagger - x_\alpha = \int (1 - g_\alpha(\lambda)) dE_\lambda x^\dagger$$

Hence, we define,

$$r_\alpha(\lambda) = 1 - \lambda g_\alpha(\lambda), \tag{5}$$

$$r_\alpha(0) = 1, \tag{6}$$

$$x^\dagger - x_\alpha = r_\alpha(T^*T)x^\dagger \tag{7}$$

Theorem (4.1)

Let for all $\alpha > 0$*, and an* $\varepsilon > 0$, $g_\alpha : [0, \|T\|^2] \to R$ *fulfills the following assumptions*

- g_α *is piece wise continuous.*
- *There is a* $C > 0$ *such that*

$$|\lambda g_\alpha(\lambda)| \leqslant C \tag{8}$$

- *and,*

$$\lim_{\alpha \to 0} g_\alpha(\lambda) = \frac{1}{\lambda} \tag{9}$$

for all $\lambda \in [0, \|T\|^2]$.

Theorem (Continued)

Then, for all $y \in D(T^\dagger)$,

$$\lim_{\alpha \to 0} g_\alpha(T^*T)T^*y = x^\dagger \tag{10}$$

and, If $y \notin D(T^\dagger)$,

$$\lim_{\alpha \to 0} \|g_\alpha(T^*T)T^*y\| = +\infty \tag{11}$$

Theorem (4.2)

Let g_α and C be as defined in theorem 4.1. For any $\alpha > 0$, let

$$G_\alpha = \sup\left\{|g_\alpha(\lambda)| \mid \lambda \in [0, \|T\|^2\right\} \tag{12}$$

Then,

$$\left\|Tx_\alpha - Tx_\alpha^\delta\right\| \leqslant C\delta \tag{13}$$

and

$$\left\|x_\alpha - x_\alpha^\delta\right\| \leqslant \delta\sqrt{CG_\alpha}. \tag{14}$$

Remark

Thus, for the total error we have the estimate

$$\left\|x_\alpha^\delta - x^\dagger\right\| \leqslant \left\|x_\alpha - x^\dagger\right\| + \delta\sqrt{CG_\alpha}$$

By Theorem 4.1, the first term in this estimate goes to zero if $y \in D(T^\dagger)$.
However, since $g_\alpha(\lambda) \to 1/\lambda$ as $\alpha \to 0$

$$\lim_{\alpha\to 0} G_\alpha = +\infty$$

Hence for any fixed $\delta > 0$,

$$\delta\sqrt{CG_\alpha} \text{ explodes} \to \left\|x_\alpha - x_\alpha^\delta\right\| \text{ explodes.}$$

Theorem (4.3)

Let g_α fulfill the assumptions of theorem 4.1. $\mu, \rho > 0$. Let $\omega_\mu : (0, \alpha_0) \to R$ be such that for all $\alpha \in (0, \alpha_0)$ and $\lambda \in [0, \|T\|^2]$,

$$\lambda^\mu |r_\alpha(\lambda)| \leqslant \omega_\mu(\alpha) \tag{15}$$

holds. Then for all $x \in X_{\mu,\rho}$,

$$\left\|x_\alpha - x^\dagger\right\| \leqslant \omega_\mu(\alpha)\rho \tag{16}$$

$$\left\|Tx_\alpha - Tx^\dagger\right\| \leqslant \omega_{\mu+\frac{1}{2}}(\alpha)\rho \tag{17}$$

Corollary (4.4)

Let the assumptions of Theorem 4.3 hold with

$$\omega_\mu(\alpha) = c\alpha^\mu \tag{18}$$

for some $c > 0$, and assume that G_α fulfills,

$$G_\alpha = O(\frac{1}{\alpha}) \quad as \quad \alpha \to 0 \tag{19}$$

Then, with parameter choice rule

$$\alpha \sim (\frac{\delta}{\rho})^{\frac{2}{2\mu+1}} \tag{20}$$

the regularization method (R_α, α) is of optimal order in $X_{\mu,\rho}$.

Example

We consider the initial value problem

$$\mu_\delta'(t) + T^*T\mu_\delta(t) = T^*y^\delta, \quad t \in R_0^+$$
$$\mu_\delta(0) = 0$$

Here, $\mu_\delta : R_0^+ \to X$.We denote by $x_\alpha^\delta = \mu_\delta(\frac{1}{\alpha})$.

Omitting δ, we get

$$v(t) = \int \gamma(t,\lambda) dE_\lambda T^* y^\delta$$
$$\gamma(t,\lambda) = \frac{1 - e^{-\lambda t}}{\lambda}$$

Showalter's Method

$$\int_0^\infty e^{-sT^*T} ds T^* y = T^\dagger y \tag{21}$$

Example

Let for $\alpha \in (0.\alpha_0), \lambda \in [0, \|T\|^2]$,

$$g_\alpha = \begin{cases} \frac{1}{\lambda}, & \lambda \geqslant \alpha \\ 0, & \lambda < \alpha \end{cases}$$

and we get,

$$x_\alpha^\delta = g_\alpha(T^*T)T^*y^\delta = \int_\alpha^{\|T\|^2} \frac{1}{\lambda} dE_\lambda T^* y^\delta$$

This is the truncated singular value expansion.

Theorem (4.5)

Assume that $T : L^2(I) \to Y$ is bounded, where I is a compact interval of R, with

$$R(T^*) \subseteq C(I)$$

where $C(I)$ is the space of continuous functions on I with supremum norm, and that

$$(x)^\dagger \in R(T^*).$$

Then x_α converges to $x^\dagger$ in $C(I)$, i.e. uniformly on I.

Table of Content

In the previous section we have seen that for regularization methods for which (18) holds,

$$\left\| x_\alpha - x^\dagger \right\| = O(\alpha^\mu) \tag{22}$$

if

$$x^\dagger \in X_\mu \tag{23}$$

The statement that (23) is not only sufficient but also necessary for (22) is called a converse result.

The term saturation is used to describe the behavior of some regularization methods for which,

$$\left\|x_\alpha - x^\dagger\right\| = O(\delta^{\frac{2\mu}{2\mu+1}}) \tag{24}$$

does not hold for all $\mu > 0$, but only upto a finite value μ_0, called the "Qualification" of the the method.
Equivalently, μ_0 is the largest values such that

$$\lambda^\mu |r_\alpha(\lambda)| = O(\alpha^\mu)$$

holds for all $0 < \mu \leqslant \mu_0$

Theorem (4.6)

Let x_α be as defined earlier and g_α fulfills the assumptions of Theorem 4.1. Assume that μ is such that

$$\lambda^\mu |r_\alpha(\lambda)| \geqslant \gamma\alpha^\mu$$
$$\lambda \in [c\alpha, \|T\|^2]$$

Then $\left\| x_\alpha - x^\dagger \right\| = O(\alpha^\mu)$ *implies* $x^\dagger \in X_\mu$

Lemma (4.7)

If,

$$\left\|E_t x^\dagger\right\|^2 = \int_0^t 1 d\left\|E_\lambda x^\dagger\right\|^2 = O(t^{2\mu}) \tag{25}$$

holds, then

$$x^\dagger \in \bigcup_{v<\mu} X_v \tag{26}$$

Proposition (4.8)

let g_α fulfill the assumptions of Theorem 4.1. Assume,

$$G_\alpha \leqslant \frac{\hat{c}}{\alpha}, \quad \alpha > 0 \tag{27}$$

holds with a suitable constant $\hat{c} > 0$,then $\left\|x_\alpha - x^\dagger\right\| = O(\alpha^\mu)$ implies (25) *and* (26).

Table of Content

The "Discrepancy Rule" is due to Marozov, and is a a-posteriori parameter choice rule.
Let g_α and r_α be as defined earlier.Furthermore, let

$$\tau > sup\{|r_\alpha(\lambda)| \mid \alpha > 0, \lambda \in [0, \|T\|^2]\} \tag{28}$$

The regularization parameter defined via the *discrepancy Principle* is

$$\alpha(\delta, y^\delta) = sup\{\alpha > 0 \mid \|Tx_\alpha^\delta - y^\delta\| \leqslant \tau\delta\} \tag{29}$$

Intuition

We want to solve $Tx = y$, but instead of y we only have noisy data y^δ and we know that $\|y - y^\delta\| \leqslant \delta$.

Thus it doesnot make sense to ask for an approximate solution $\hat{x}$ with the discrepancy $\|T\hat{x} - y^\delta\| < \delta$. A residual of the order of δ is the best we can ask for.

Assumptions

- $y \in R(T)$
- $\mu_0 > \frac{1}{2}$
- $\omega_\mu \sim \alpha^\mu \quad \mathit{for} \quad 0 < \mu \leqslant \mu_0$

Theorem (4.9)

The regularization method (R_α, α)*, where* α *is defined via the discrepancy principle, is convergent for all* $y \in R(T)$*, and of optimal order in* $X_{\mu,\rho}$*, for* $\mu \in (0, \mu_0 - 1/2]$

Proposition (4.10)

Let K be compact, $R_\alpha = (K^*K + \alpha I)^{-1}K^*$,$\alpha$ be defined by the discrepancy rule.If

$$\|x_\alpha^\delta - x^\dagger\| = o(\sqrt{\delta}) \tag{30}$$

holds for all $y \in R(T)$ and $y^\delta \in Y$, then $R(K)$ is finite dimensional.