Basic information: Parameter choice and iterative regularization methods

Even though there are theoretic results for a suitable a-priori parameter choice rule via $\alpha \approx \left(\frac{\delta}{\rho}\right)^{\frac{2}{2\mu+1}}$ for $x^\dagger \in X_{\mu,\rho}$ using spectral filtering, since $x^\dagger$ is unknown, also μ and ρ are, making this parameter choice infeasible. To remedy this, one can consider a-posteriori parameter choice rules:

The discrepancy principle works as follows: Choose some $\tau > \sup\left\{|r_\alpha(\lambda)| \mid \alpha \in]0, \alpha_0[,\ \lambda \in [0, \|T\|^2]\right\}$. For $M_{\tau,\delta} := \{\tilde{\alpha} > 0 \mid \|Tx^\delta_{\tilde{\alpha}} - y^\delta\| \leq \tau\delta\}$, choose $\alpha(\delta, y^\delta) = \sup M_{\tau,\delta}$ which is possible if $y \in \mathrm{Rg}(T)$. In fact, choosing α not the supremum, but $\alpha \in M_{\tau,\delta}$ such that there is $k \in]1, K]$ (some fixed $K > 1$) with $k\alpha \notin M_{\tau,\delta}$ is more manageable (while maintaining its theoretical properties). Note that using an (up to μ_0) order optimal linear regularization together with the discrepancy method yields order-optimality up to $\mu_0 - \frac{1}{2}$.

The improved a-posteriori rule consists in solving the equation $f(\alpha, y^\delta) = c\delta^2$ for α, where $f(\alpha, w) = 2\langle F_\alpha(TT^*)Qw, Qw\rangle$, with $F_\alpha(\lambda) := \left(\frac{\partial G_\alpha}{\partial \alpha}(\lambda)\right)^{-1} \frac{\partial g_\alpha}{\partial \alpha}(\lambda) r_\alpha(\lambda)$, the orthogonal projection Q onto the closure of the range of T, and $c = \sup_{\lambda,\alpha} |\lambda g_\alpha(\lambda)|$. Naturally, we must assume $\alpha \mapsto G_\alpha$ and $\alpha \mapsto g_\alpha(\lambda)$ to be continuously differentiable (w.r.t. α, for each λ) and such that $\frac{\partial g_\alpha}{\partial \alpha} \left(\frac{\partial G_\alpha}{\partial \alpha}\right)^{-1}$ is bounded (for all α, λ). Moreover, we assume $F_\alpha(\lambda)$ is strictly increasing w.r.t. α (for all λ) and $Qy^\delta \neq 0$. One weakens this to solving $f(\alpha, y^\delta) = \tau\delta^2$ for some $\tau > 0$, in fact there is a unique solution α if $\tau \in]0, h(y^\delta)\delta^{-2}[$ with $h(w) = 2\int_{\sigma(TT^*)} \lim_{\alpha\to\infty} F_\alpha(\lambda)\, \mathrm{d}\|E_\lambda Qw\|^2$ (E the spectral measure of TT^*). Setting $L := 2\sup_{\alpha,\lambda} F_\alpha(\lambda)$ and $\tau \in]L, \frac{h(y^\delta)}{\delta^2}[$, the improved a-posteriori choice yields a quasi-optimal estimate, i.e., this choice yields $\|x^\delta_\alpha - x^\dagger\| \leq \nu \inf\{\|x_{\tilde{\alpha}} - x^\dagger\| + \delta\sqrt{cG_{\tilde{\alpha}}} \mid \tilde{\alpha} > 0\}$, for some constant $\nu > 0$.

Similarly to the matrix-free approaches discussed in the previous sheet, one might want to consider iterative regularization methods, often representing cut-off power series. It is convenient if these iterative methods can be again understood as spectral filtering (understanding $\frac{1}{\alpha}$ as the stopping time), since then previously discussed techniques for determining properties such as regularization or order optimality are applicable. One such regularization method is the Landweber iteration, which for $x_0 = 0$ consists in $x_{k+1} = x_k + wT^*(y^\delta - Tx_k)$ for some $w > 0$, $w \leq \frac{1}{\|T\|^2}$ and yields a regularization without saturation.

Example 5.1) [Chebyshev regularization]

We consider $T \in L(X, Y)$ between Hilbert spaces with $\|T\| \leq 1$. For $y^\delta \in Y$ we consider the iteration

$$\begin{cases} x_0 = 0, \qquad x_1 = x_0 + \frac{4}{3}T^*(y^\delta - Tx_0), \\ x_k = 2\frac{2k-1}{2k+1}x_{k-1} - \frac{2k-3}{2k+1}x_{k-2} + 4\frac{2k-1}{2k+1}T^*(y^\delta - Tx_{k-1}) \qquad \text{for } k = 2, 3, \dots. \end{cases} \tag{1}$$

a) Show that $x_k = g_k(T^*T)T^*y^\delta$ for some polynomials g_k and conclude their recursive construction.

b) Show that the residual $r_k := 1 - \lambda g_k$ for $k \geq 0$ is again a polynomial that satisfies $r_k = (-1)^k \frac{T_{2k+1}(\sqrt{\lambda})}{(2k+1)\sqrt{\lambda}}$, where $T_m(\lambda) := \cos(m \arccos(\lambda)) = 2\lambda T_{m-1}(\lambda) - T_{m-2}(\lambda)$ the Chebyshev polynomials (it is easy to see that $T_0(\lambda) = 1$, $T_1(\lambda) = \lambda$).

c) When setting $R_\alpha = g_{\lfloor\frac{1}{\alpha}\rfloor}(T^*T)T^*$, show that $(R_\alpha)_\alpha$ is a continuous regularization.

Remark. *This illustrates an iterative reconstruction method, which, unlike the Landweber iteration is not represented by a cut-off power series, and in fact is not given in a convenient explicit form, nonetheless, previously discussed techniques can be applied using induction and recursive properties.*

Example 5.2) [Improved a-posteriori choice for Showalter]
We again consider the Showalter regularization induced by $g_\alpha(\lambda) = \frac{1-e^{-\frac{\lambda}{\alpha}}}{\lambda}$ (see Example 3.3), and want to apply the improved a-posteriori choice.

a) Compute G_α, $\partial_\alpha g_\alpha$ and $\partial_\alpha G_\alpha$. Conclude the equivalence

$$f(\alpha, y^\delta) = \tau\delta^2 \quad \Leftrightarrow \quad \|\exp\left(-\frac{TT^*}{\alpha}\right) Qy^\delta\|^2 = \tau\delta^2 \quad \Leftrightarrow \quad \|Tx_\delta^\alpha - Qy^\delta\| = \sqrt{\frac{\tau}{2}}\delta. \quad (2)$$

Hint: for the computation of G_α, show that $\lambda \mapsto g_\alpha(\lambda)$ is monotone, $1 + z \leq e^z$ for $z \geq 0$.

b) Show that for $\tau \in]0, \frac{2\|Qy^\delta\|^2}{\delta^2}[$ there is a unique solution of (2). Further, show that when choosing $\alpha(\delta, y^\delta)$ according to (2) for some $\tau \in]2, 2\frac{\|Qy^\delta\|^2}{\delta^2}[$, then $(R_\alpha, \alpha(\delta, y^\delta))$ yields an order optimal regularization method for any $\mu > 0$. **Hint:** First, show that R_α with the a-priori rule is order optimal.

Remark. *Note that, as the name suggests, these types of parameter choices potentially require many evaluations of R_α for different α, but $f(\alpha, y^\delta)$ being monotone invites simple solution algorithms. Also note the similarity of (2) to the discrepancy principle.*

Example 5.3) [Implementation of Landweber with discrepancy principle]
We consider the Landweber regularization to a linear problem $Tx = y$.

- Explain how the discrepancy principle for such a method can be implemented practically, and determine suitable τ. Why is the improved a-posteriori rule not feasible for iterative methods?
- Implement a Matlab function '[rec,iter]=Landweber(A, ν, y^δ, δ, τ)', which given a matrix A with estimate of the norm ν, and data y^δ with noise level δ applies the Landweber method with discrepancy principle (with τ), and returns the reconstruction as well, as the amount of iterations applied.
- Test your program with the dataset walnut data[1] with Data328.mat (see documentation) with $\delta \in \{3, 9, 15\}$ and $\tau = 1.1$ and visualize the solutions and the number of iterations. You may use the poweriteration.m file to estimate the norm of the operator (10 iterations or so should suffice, look whether the value stabilizes).

[1] Dataset courtesy of Keijo Hämäläienen et al, Tomographic X-ray data of a walnut. ArXiv eprint 1502.04064, 2015.