# Minimization strategy for choice of the regularization parameter in case of roughly given or unknown noise level

Uno Hämarik, Reimo Palm, Toomas Raus

Vienna, July 23, 2009

# The problem

- We consider an operator equation

$$Au = f_0,$$

  where $A$ is a linear bounded operator between real Hilbert spaces, $f_0 \in \mathcal{R}(A) \Rightarrow \exists$ solution $u_* \in H$.
- Instead of exact data $f_0$, noisy data $f$ are available.
- Knowledge of $\|f_0 - f\|$:
  - Case 1: exact noise level $\delta$: $\|f_0 - f\| \leq \delta$
  - Case 2: approximate noise level $\delta$: $\lim \|f_0 - f\|/\delta \leq C$ as $\delta \to 0$
  - Case 3: no information about $\|f_0 - f\|$

# Methods

Let $D = A^*$ in case $A \neq A^*$ and $D = I$ in case $A = A^* \geq 0$.

- (Iterated) Tikhonov method ($D = A^*$) and (iterated) Lavrentiev method ($D = I$): $u_\alpha = u_{m,\alpha}$, where $u_{k,\alpha} = (\alpha I + DA)^{-1}(\alpha u_{k-1,\alpha} + Df)$, $k = 1, \ldots, m$.
- Landweber method: $u_n = u_{n-1} - \mu D(Au_{n-1} - f)$, $\mu \in (0, 1/\|DA\|)$, $n = 1, 2, \ldots$
  We realized Landweber method by operator iterations

$$u_n = (I - DAC_k)u_0 + C_k Df, \qquad n = m^k,\ k = 1, 2, \ldots,$$

$$C_k = C_{k-1} \sum_{j=0}^{m-1} (I - DAC_{k-1})^j, \qquad C_0 = \mu I, \quad k = 1, 2, \ldots$$

# Rules for choice of $\alpha$ in (iterated) Tikhonov method

- Case 1: $\delta$ with $\|f - f_0\| \leq \delta$. **Rule D**: $\|r_{m,\alpha_D}\| = \delta$, where $r_{m,\alpha} = Au_{m,\alpha} - f$. **Rule MEe**: 1) find by the monotone error rule $\alpha_{ME}$ as solution of equation $(r_{m,\alpha}, r_{m+1,\alpha})/\|r_{m+1,\alpha}\| = \delta$; 2) take $\alpha_{MEe} = \min(0.53\alpha_{ME}, 0.6\alpha_{ME}^{1.06})$.
- Case 2: approximate $\delta$ is known with $\|f - f_0\|/\delta \leq$ const ($\delta \to 0$). **Rule DM**: 1) find $\underline{\alpha}$ as maximal solution of equation $\sqrt{\alpha}\|u_{m,\alpha} - u_{m+1,\alpha}\| = b(m+1)^{m+1}/m^m \cdot \delta$, $b =$ const; 2) fix $s \in (0,1)$, $q \in (0,1)$ and find $\alpha(\delta) = \operatorname{argmin}\{\Phi(\alpha)/\alpha^{s/2},\ \alpha \in [\underline{\alpha}, 0.4m + 0.6]\}$, where $\Phi(\alpha) = \|u_\alpha - u_{q\alpha}\|(1 + \frac{\alpha}{\|A\|^2})^{1/(2m)}$. In computations we used $b = s = 0.01$, $q = 0.9$ and if the first equation had no solution, then the largest local minimum of $\sqrt{\alpha}\|u_{m,\alpha} - u_{m+1,\alpha}\|$ was taken as $\underline{\alpha}$.

# Rules for Case 3 in (iterated) Tikhonov method

$m{+}2,\alpha)^{1/2}$

- **Hanke-Raus rule**: $\alpha_{HR}$ is the global minimizer of the function $\varphi$
  $_{HR}(\alpha) = d_{MD}(\alpha)/\sqrt{\alpha}$, where $d_{MD}$ $m{+}1,\alpha - u$

interval $[m\lambda$ is the minimizer of $\varphi$ $(\alpha) = \|u_{m,\alpha} - u_{2m,\alpha}\|$ on the

$_{min}, 1]$, where $\lambda$ $A$ is the smallest eigenvalue of $m{+}1,\alpha, u$

$^{*}_{l}$ ere $\alpha$

$m,\alpha - u$

$(u$ ,

$\kappa(\alpha) = 1 +$

$(\alpha)/\sqrt{\alpha}$ is 3 and 5 times larger than its value at the minimum

- **Rules QO1**, **D1**, **R21** and **DR21** choose the parameter by the largest local minimum in functions $\varphi_{\mathsf{QN}}(\alpha)\alpha^{1/3}$, $\|r_{m,\alpha}\|\alpha^{-1/6}$, $d_{\mathsf{R2}}(\alpha)\alpha^{-1/6}$ and $\sqrt{\|r_{m,\alpha}\|d_{\mathsf{R2}}(\alpha)}\,\alpha^{-1/6}$.
- **Rule HR2** chooses the parameter as the global minimizer of $\varphi_{\mathsf{HR2},\tau}(\alpha) = d_{\mathsf{HR}}(\alpha)^{1-\left(\frac{d_{\mathsf{R2}}(\alpha)}{d_{\mathsf{MD}}(\alpha)}\right)^{\tau}} d_{\mathsf{R2}}(\alpha)^{\left(\frac{d_{\mathsf{R2}}(\alpha)}{d_{\mathsf{MD}}(\alpha)}\right)^{\tau}} / \sqrt{\alpha}$. with $\tau = 0.15$.
- **Rule QHR2** chooses local minimizer of the function $\varphi_{\mathsf{QN}}(\alpha)\kappa(\alpha)$ for which the function $\varphi_{\mathsf{HR2},\tau}(\alpha)$ with $\tau = 0.04$ is minimal.
- **Rule QOHR** chooses local minimizer of the function $\varphi_{\mathsf{QN}}(\alpha)\kappa(\alpha)$ for which the function $d_{\mathsf{MD}}(\alpha)/\sqrt{\alpha}$ is minimal.

# Rules for choice of $\alpha$ in (iterated) Lavrentiev method

- Case 1: $\delta$ with $\|f - f_0\| \le \delta$. **Rules MD**, **MEa** choose $\alpha$'s from equations $\|r_{m+1,\alpha}\| = 1.143\delta$ and $\frac{\|r_{m+1,\alpha}\|^2}{\|r_{m+2,\alpha}\|} = 1.364\delta$, respectively. **Rules MEn**, **MEl** choose $\alpha$ and $\alpha_i = 1.2^{-i}$ from equations $\frac{(r_{2m+1,\alpha}, r_{1,0.2\alpha})}{(r_{2m+2,\alpha}, r_{2,0.2\alpha})^{1/2}} = 1.096\delta$, and $\frac{(u_{m,\alpha_i} - u_{m,\alpha_{i-1}}, u_{m,\alpha_{5+i}} - u_{m,\alpha_{5+i-1}})}{\|u_{m+1,\alpha_{5+i}}, u_{m+1,\alpha_{4+i}}\|} \cdot \frac{\alpha_i}{q-1} \cdot \frac{m+1}{m^2} \approx 1.004\delta$,, respectively.
- Case 2: approximate $\delta$ is known with $\|f - f_0\|/\delta \le$ const $(\delta \to 0)$. **Rule DM**: 1) find $\underline{\alpha}$ as minimal solution of equation $(r_{\alpha,1}, Ar_{\alpha,2})/\sqrt{\alpha} = \frac{1}{\sqrt{2m+3}} \cdot \delta$; 2) fix $s \in (0,1)$, $q \in (0,1)$ and find $\alpha(\delta) = \operatorname{argmin}\{\Phi(\alpha)/\alpha^s,\ \alpha \in [\underline{\alpha}, m]\}$, where $\Phi(\alpha) = \|r_{\alpha,1}\| \cdot (r_{\alpha,1}, r_{\alpha,2})^{1/2}/(r_{\alpha,2}, r_{\alpha,3})^{1/2}/\alpha$ in case $m = 1$, and $\Phi(\alpha) = \|u_\alpha - u_{q\alpha}\|(1 + \frac{\alpha}{\|A\|})^{0.005}$ in case $m \ge 2$.

# Case 3 for (iterated) Lavrentiev method

- **Rules QOC**, **QOmC** minimize the functions $\|u_\alpha - u_{q\alpha}\|(1+\frac{\alpha}{\|A\|})^{0.005}$ and $\|r_{m,\alpha}\| \cdot (r_{m,\alpha}, r_{m+1,\alpha})^{1/2}/(r_{m+1,\alpha}, r_{m+2,\alpha})^{1/2}$ on the interval $[\underline{\alpha}, m]$, where $\underline{\alpha}$ is the largest $\alpha$, for which the values of these functions are 1.4 times larger than their values at the minimum point.

# Convergence and error estimate for Rule DM (Case 2)

Let the parameter $r(\delta) = \alpha^{-1}$ in $\gamma = m$ times iterated Tikhonov or Lavrentiev method or $r(\delta) = n$ in Landweber method with $\gamma = \mu$ be chosen according to Rule DM. If $\lim \|f_0 - f\|/\delta \leq C$ as $\delta \to 0$, then $\|u_r - u_*\| \to 0$ and in case $b \geq 1$ the following error estimates hold:

1. If $\|f_0 - f\| \leq \max\{\delta, \tilde{\delta}(r(\delta))\}$, where $\tilde{\delta}(r(\delta)) := \frac{1}{2}\|r_{m+1,\alpha(\delta)}\|$ in (iterated) Tikhonov or Lavrentiev method, $\tilde{\delta}(r(\delta)) = \frac{1}{2}\|Au_{n(\delta)} - f\|$ in Landweber method, then
$$\|u_{r(\delta)} - u_*\| \leq C' \frac{1}{1 - s/\sigma} \inf_{r \geq 0} \Psi(r),\ \Psi(r) = e_r^0 + \gamma r \max\{\delta, \|f_0 - f\|\},$$
$e_r^0 = \|u_r^0 - u_*\|$, $u_r^0$ is the approximation at the exact data.
2. If $\max\{\delta, \tilde{\delta}(r(\delta))\} < \|f_0 - f\| \leq \frac{1}{2}\tilde{\delta}(1)$, then
$$\|u_{r(\delta)} - u_*\| \leq C''(\|f_0 - f\|/\tilde{\delta}(r(\delta)))^{\sigma/s} \inf_{r \geq 0} \Psi(r).$$

## Test problems

- 10 test problems of P. C. Hansen. Typically integral equations of the first kind, arising from various applications: inverse heat equation, gravity surveying, inverse Laplace transform etc.
- Supposable noise levels: $\delta = 0.5,\ 10^{-1},\ 10^{-2},\ \dots,\ 10^{-6}$. Actual noise level is $d\delta$ where $d = 1,\ 10$. Each problem was solved 10 times.
- The discretized problems (discretization parameter = 100) were solved using different parameter choice rules.
- In the following tables we present averages of error ratios $\|u_r - u_*\|/e_{\text{opt}}$ (over $\delta$'s and 10 runs) in case $\delta = d\|f_\delta - f\|$, where $e_{\text{opt}} = \min\{\|u_r - u_*\| : r \geq 0\}$ and $d = 1,\ 10$ ($d > 1$ corresponds to overestimation of noise level); $e_{\text{D}}$ and $e_{\text{D},10}$ correspond to $\alpha_{\text{D}}$ in cases $d = 1$ and $d = 10$ etc.

## Results for Tikhonov method

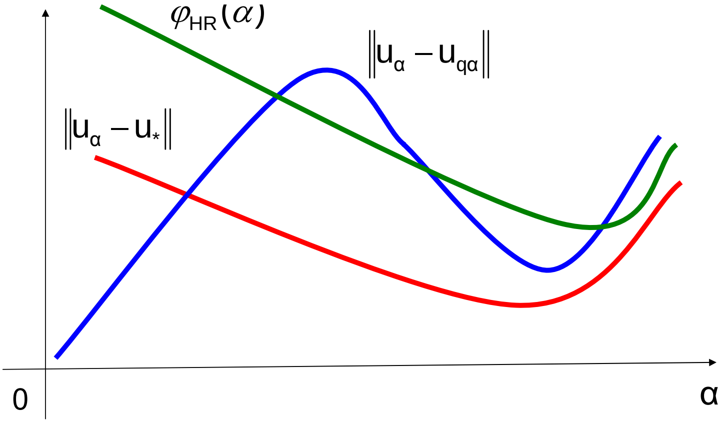
| Problem | $e_{\mathrm{D}}$ | $e_{\mathrm{D},10}$ | $e_{\mathrm{MEe}}$ | $e_{\mathrm{MEe},10}$ | $e_{\mathrm{DM}}$ | $e_{\mathrm{DM},10}$ | $e_{\mathrm{HR}}$ | $e_{\mathrm{QN}}$ | $e_{\mathrm{QOC}}$ |
|---|---|---|---|---|---|---|---|---|---|
| baart | 1.41 | 3.21 | 1.38 | 2.46 | 1.72 | 2.60 | 2.76 | 1.56 | 1.89 |
| deriv2 | 1.20 | 3.44 | 1.05 | 1.97 | 1.14 | 1.50 | 1e3 | 1.87 | 1.14 |
| foxgood | 1.38 | 11.2 | 1.37 | 6.98 | 2.09 | 5.12 | 8.16 | 2.18 | 2.18 |
| gravity | 1.15 | 4.42 | 1.09 | 2.71 | 1.11 | 2.08 | 2.81 | 1.13 | 1.13 |
| heat | 1.06 | 2.76 | 1.03 | 1.82 | 1.17 | 1.36 | 1.70 | 1e5 | 1.25 |
| ilaplace | 1.26 | 2.58 | 1.17 | 2.00 | 1.21 | 1.67 | 2.05 | 1.19 | 1.19 |
| phillips | 1.03 | 4.20 | 1.04 | 2.43 | 1.09 | 1.78 | 1e5 | 1.08 | 1.08 |
| shaw | 1.31 | 3.20 | 1.26 | 2.43 | 1.48 | 2.25 | 2.57 | 1.44 | 1.47 |
| spikes | 1.02 | 1.08 | 1.02 | 1.06 | 1.04 | 1.07 | 1.07 | 1.04 | 1.05 |
| wing | 1.19 | 1.50 | 1.18 | 1.47 | 1.48 | 1.55 | 1.56 | 1.43 | 1.79 |
| Mean | 1.20 | 3.75 | 1.16 | 2.53 | 1.35 | 2.10 | 1e4 | 1e4 | 1.42 |

## Tikhonov method: averages of error ratios.

| Problem | $e_{QOC}$ | $e_{R2C}$ | $e_{R2LC}$ | $e_{QO1}$ | $e_{D1}$ | $e_{R21}$ | $e_{DR21}$ | $e_{HR2}$ | $e_{QHR2}$ | $e_{QOHR}$ |
|---|---|---|---|---|---|---|---|---|---|---|
| baart | 1.89 | 1.84 | 1.52 | 2.29 | 2.00 | 3.68 | 2.05 | 2.46 | 1.75 | 2.79 |
| deriv2 | 1.14 | 1.10 | 1.16 | 1.06 | 1.65 | 1.07 | 1.24 | 1.19 | 1.21 | 1.21 |
| foxgood | 2.18 | 2.11 | 2.11 | 1.80 | 3.52 | 1.89 | 2.25 | 3.26 | 2.08 | 9.56 |
| gravity | 1.13 | 1.11 | 1.11 | 1.40 | 1.45 | 1.43 | 1.13 | 1.28 | 1.10 | 1.10 |
| heat | 1.25 | 1.19 | 1.27 | 1.13 | 1.22 | 1.18 | 1.18 | 1.19 | 1.16 | 1.82 |
| i_laplace | 1.19 | 1.18 | 1.18 | 1.20 | 1.48 | 1.21 | 1.13 | 1.47 | 1.23 | 1.27 |
| phillips | 1.08 | 1.08 | 1.07 | 1.16 | 1.21 | 1.27 | 1.16 | 1.09 | 1.09 | 1.09 |
| shaw | 1.47 | 1.46 | 1.45 | 1.73 | 1.71 | 1.80 | 1.45 | 1.99 | 1.48 | 2.38 |
| spikes | 1.05 | 1.05 | 1.04 | 1.08 | 1.04 | 1.07 | 1.06 | 1.07 | 1.05 | 1.06 |
| wing | 1.79 | 1.44 | 1.42 | 1.83 | 1.48 | 1.81 | 1.80 | 1.49 | 1.47 | 1.55 |
| Mean | 1.42 | 1.35 | 1.33 | 1.47 | 1.67 | 1.64 | 1.44 | 1.65 | 1.36 | 2.38 |

## Behaviour of quasioptimality criterion and Hanke-Raus rule

Quasioptimality criterion minimizes in Tikhonov method $\|u_\alpha - u_{q\alpha}\|$, Hanke-Raus rule minimizes $\varphi_{\mathrm{HR}}(\alpha)$.



## Lavrentiev method: averages of error ratios.

| Problem | $e_{\mathrm{MD}}$ | $e_{\mathrm{MD},10}$ | $e_{\mathrm{MEa}}$ | $e_{\mathrm{MEa},10}$ | $e_{\mathrm{MEn}}$ | $e_{\mathrm{MEn},10}$ | $e_{\mathrm{MEl}}$ | $e_{\mathrm{MEl},10}$ | $e_{\mathrm{QOC}}$ | $e_{\mathrm{QOmC}}$ |
|---|---|---|---|---|---|---|---|---|---|---|
| deriv2 | 1.79 | 3.84 | 1.24 | 3.18 | 1.20 | 3.18 | 1.05 | 2.63 | 1.11 | 1.11 |
| foxgood | 1.03 | 2.15 | 1.02 | 1.61 | 1.00 | 1.88 | 1.00 | 1.80 | 1.05 | 1.00 |
| gravity | 1.02 | 2.15 | 1.02 | 1.62 | 1.00 | 1.88 | 1.00 | 1.81 | 1.06 | 1.00 |
| phillips | 1.01 | 2.34 | 1.01 | 1.64 | 1.01 | 1.97 | 1.01 | 1.89 | 1.06 | 1.00 |
| shaw | 1.01 | 1.70 | 1.01 | 1.48 | 1.00 | 1.57 | 1.00 | 1.52 | 1.08 | 1.03 |
| Mean | 1.17 | 2.43 | 1.06 | 1.90 | 1.04 | 2.10 | 1.01 | 1.93 | 1.07 | 1.03 |

# Landweber method: averages of error ratios

| Problem | $e_{D}$ | $e_{D,10}$ | $e_{De}$ | $e_{De,10}$ | $e_{HR}$ | $e_{HRmC}$ | $e_{Neub}$ | $e_{NeubmC}$ |
|---|---|---|---|---|---|---|---|---|
| baart | 1.47 | 3.67 | 1.40 | 3.03 | 2.71 | 2.50 | 1.80 | 1.75 |
| deriv2 | 1.33 | 4.02 | 1.05 | 3.02 | 1e+3 | 1.95 | 1e+3 | 1.59 |
| foxgood | 2.35 | 21.45 | 1.76 | 8.94 | 7.75 | 2.98 | 5.01 | 3.63 |
| gravity | 1.44 | 6.94 | 1.16 | 3.70 | 2.70 | 1.83 | 1.27 | 1.61 |
| heat | 1.22 | 3.60 | 1.05 | 2.83 | 1.67 | 7.39 | 1e+4 | 2.37 |
| ilaplace | 1.35 | 3.17 | 1.24 | 2.27 | 2.00 | 1.39 | 1.24 | 1.30 |
| phillips | 1.37 | 7.83 | 1.08 | 3.78 | 2e+5 | 1.44 | 2e+5 | 1.49 |
| shaw | 1.40 | 3.83 | 1.29 | 2.70 | 2.58 | 2.11 | 1.60 | 1.56 |
| spikes | 1.02 | 1.09 | 1.02 | 1.07 | 1.07 | 1.07 | 1.05 | 1.04 |
| wing | 1.21 | 1.62 | 1.18 | 1.48 | 1.55 | 1.41 | 1.48 | 1.41 |
| Mean | 1.42 | 5.72 | 1.22 | 3.28 | 2e+4 | 2.41 | 2e+4 | 1.78 |

## Comparison of methods: means of minimal relative error for $\delta = 10^{-4}$.

| Problem | Tikh | Lavr | Landw | CGLS | CGME |
|---|---|---|---|---|---|
| baart | 6.27e-2 | – | **6.20e-2** | 8.63e-2 | 1.16e-1 |
| deriv2 | **1.07e-1** | 1.23e-1 | **1.07e-1** | 1.09e-1 | 1.29e-1 |
| foxgood | 4.95e-3 | 2.60e-2 | **4.51e-3** | 5.55e-3 | 8.27e-3 |
| gravity | 7.12e-3 | 2.15e-2 | **6.80e-3** | 6.99e-3 | 1.48e-2 |
| heat | 1.80e-2 | – | **1.70e-2** | **1.70e-2** | 2.09e-2 |
| i_laplace | 7.07e-2 | – | **6.97e-2** | 7.08e-2 | 9.63e-2 |
| phillips | 5.12e-3 | 1.71e-2 | **4.69e-3** | 4.73e-3 | 8.44e-3 |
| shaw | 3.11e-2 | 6.15e-2 | **3.09e-2** | 3.56e-2 | 4.74e-2 |
| spikes | **7.88e-1** | – | 7.90e-1 | 7.98e-1 | 8.23e-1 |
| wing | **3.64e-1** | – | 3.80e-1 | 4.45e-1 | 5.95e-1 |