§2 CLASSICAL REGULARIZATION METHODS

let $\mathcal{X}, \mathcal{Y}$ be seperable Hilbert spaces (of finite or infinite dimension) and $\mathcal{K}: \mathcal{X} \to \mathcal{Y}$ be a compact operator

let $x \in \mathcal{X}$ and $y \in \mathcal{Y}$; we consider inverse problem of finding $x \in \mathcal{X}$ satisfying operator equation

$$\mathcal{K}x = y \qquad (*)$$

for given $y \in \mathcal{Y}$ (linear inverse problem)

$(*)$ is a Fredholm equation of first kind

example: Fredholm first kind integral equation of convolution type in one space dimension:

$$\mathcal{K}f(x) := \int_0^1 k(x-\tilde{x}) f(\tilde{x})\, d\tilde{x} = y(x)$$

with kernel $k(x) = c \exp\left(-\frac{1}{2\sigma^2} x^2\right)$, $c, \sigma \in \mathbb{R}$

smoothing operator

direct problem: given source f and kernel k, determine y

inverse problem: given kernel k, and observation y, determine source f

notation and analytical tools

- $\mathcal{X}$: real Hilbert space with inner product $\langle \cdot, \cdot \rangle_{\mathcal{X}}$
- domain: $\mathrm{dom}(\mathcal{K}) := \mathcal{X}$
- range: $\mathrm{range}(\mathcal{K}) := \{\mathcal{K}x \in \mathcal{Y} : x \in \mathcal{X}\}$
- kernel: $\ker(\mathcal{K}) := \{x \in \mathcal{X} : \mathcal{K}x = 0\}$
- orthogonal complement:
$$\mathcal{U}^\perp := \{x \in \mathcal{X} : \langle x, u \rangle_{\mathcal{X}} = 0 \text{ for all } u \in \mathcal{U}\}$$
- sum: $\mathcal{U} \oplus \mathcal{V} = \{u + v : u \in \mathcal{U}, v \in \mathcal{V}\}$

- $\mathcal{K}$ is continuous if $\mathcal{K}x_n \to \mathcal{K}x$ whenever $x_n \to x$
  symbol $\to$ denotes strong convergence, i.e., $x_n \to x$ means $\lim_{n\to\infty} \|x_n - x\| = 0$
- a linear operator is bounded iff induced operator norm $\|\mathcal{K}\|_{\mathcal{L}(\mathcal{X},\mathcal{Y})}$ is bounded, i.e., $\mathcal{K} \in \mathcal{L}(\mathcal{X},\mathcal{Y})$ with
  $$\|\mathcal{K}\|_{\mathcal{L}(\mathcal{X},\mathcal{Y})} := \sup_{\|x\|_{\mathcal{X}}=1} \|\mathcal{K}x\|_{\mathcal{Y}} < \infty$$
  space of bounded linear operators from $\mathcal{X}$ to $\mathcal{Y}$
- adjoint of bounded linear operator $\mathcal{K}^* \in \mathcal{L}(\mathcal{Y},\mathcal{X})$ is characterized by $\langle \mathcal{K}x, y\rangle_{\mathcal{Y}} = \langle x, \mathcal{K}^* y\rangle_{\mathcal{X}}$, $x \in \mathcal{X}$, $y \in \mathcal{Y}$,
- self-adjoint operator: $\mathcal{K} = \mathcal{K}^*$ (requires $\mathcal{X} = \mathcal{Y} = \mathcal{H}$)
- positive definiteness
  $$\lambda_{min}(\mathcal{K}) := \inf_{\|u\|_{\mathcal{H}}=1} \langle \mathcal{K}u, u\rangle_{\mathcal{H}}$$
  $$\lambda_{max}(\mathcal{K}) := \sup_{\|u\|_{\mathcal{H}}=1} \langle \mathcal{K}u, u\rangle_{\mathcal{H}}$$

  $\lambda_{min}(\mathcal{K}) \geq 0$ $\quad \mathcal{K}$ is positive semidefinite
  $\lambda_{min}(\mathcal{K}) > 0$, $u \neq 0$, $\quad \mathcal{K}$ is positive definite

let $\mathcal{X}, \mathcal{Y}$ denote normed spaces; linear operator $\mathcal{K}: \mathcal{X} \to \mathcal{Y}$ is called compact if following equivalent properties hold

- every image of any bounded subset $\mathcal{U}$ of $\mathcal{X}$ is relatively compact subset of $\mathcal{Y}$, i.e., $\overline{\mathcal{K}\mathcal{U}}$ is compact in $\mathcal{Y}$
- image in $\mathcal{Y}$ of bounded sequence $\{\varphi_n\}_{n\in\mathbb{N}} \subset \mathcal{X}$ contains convergent subsequence

well-posedness

problem (*) is well-posed if

(i) for all $y \in \mathcal{Y}$, there exists a $x \in \mathcal{X}$ with $\mathcal{K}x = y$

(ii) $\tilde{x} \neq x$ implies $\mathcal{K}\tilde{x} \neq y$ (solution is unique)

(iii) for all $\{x_n\}_{n\in\mathbb{N}}$ with $Kx_n \to y$ we have $x_n \to x$
(solution of problem depends continuously on data)

if (*) is well-posed, K has well-defined, continuous inverse K^{-1} (in particular, $\forall x\in\mathcal{X}: K^{-1}Kx = x$ and $\mathrm{range}(K) = \mathcal{Y}$)

for problem (*) to have unique solution the following conditions must be satisfied

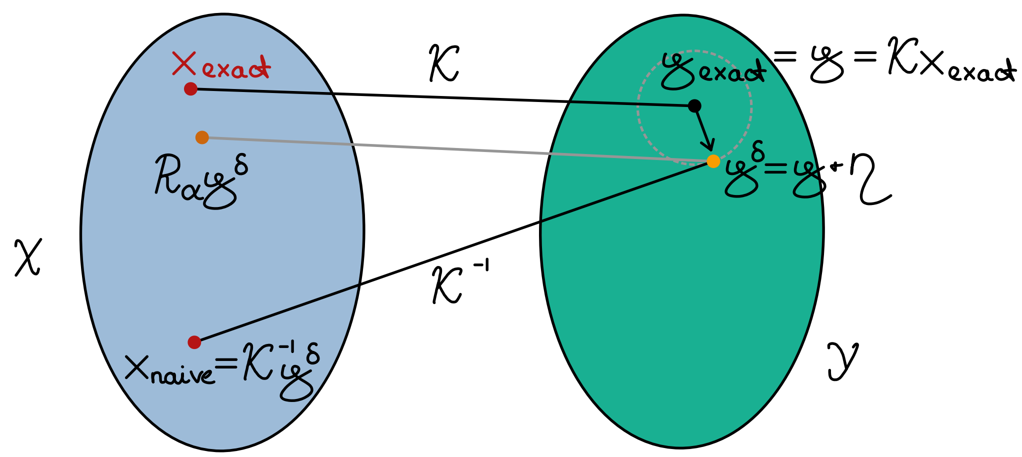
- solution exists iff $y \in \mathrm{range}(K)$ (i.e., $\mathrm{range}(K) = \mathcal{Y}$) and
- solution is unique iff $\ker(K) = \{0\}$

additional problem: y typically represents measured data, i.e., $y^\delta = y + \eta$

consequently, (*) becomes $Kx \approx y^\delta$

even if conditions (i) and (ii) guarantee existence of a unique solution of (*), they are not sufficient in practical applications

classical regularization methods are designed to overcome obstacles illustrated above



continuity requirement for K^{-1} ensures that $K^{-1}y$ and $K^{-1}y^\delta$ are close in $\mathcal{X}$ if y and y^δ are close in $\mathcal{Y}$

basic idea: instead of solving (*) exactly, we seek to find a nearby problem that
(i) is uniquely solvable
(ii) is robust/stable in sense that small errors in data do not corrupt excessively approximate solution

we will review three classical methods
(i) regularization by singular value truncation
(ii) Tikonov regularization
(iii) regularization by truncated iterative methods

truncated singular value decomposition

[proposition]

let $\mathcal{X}, \mathcal{Y}$ denote finite or infinite dimensional Hilbert spaces and $\mathcal{K}: \mathcal{X} \to \mathcal{Y}$ be as above with adjoint $\mathcal{K}^*: \mathcal{Y} \to \mathcal{X}$; then

(i) the spaces $\mathcal{X}, \mathcal{Y}$ allow for orthogonal decompositions

$$\mathcal{X} = \ker(\mathcal{K}) \oplus \ker(\mathcal{K})^\perp = \ker(\mathcal{K}) \oplus \overline{\operatorname{range}(\mathcal{K}^*)}$$

$$\mathcal{Y} = \overline{\operatorname{range}(\mathcal{K})} \oplus (\operatorname{range}(\mathcal{K}))^\perp = \overline{\operatorname{range}(\mathcal{K})} \oplus \ker(\mathcal{K}^*)$$

(ii) there exists orthonormal sets of vectors $(u_i) \in \mathcal{X}$, $(v_i) \in \mathcal{Y}$, and a sequence (σ_i) of positive numbers $\lim_{i \to \infty} \sigma_i = 0$ such that

$$\overline{\operatorname{range}(\mathcal{K})} = \overline{\operatorname{span}}\{v_i \in \mathcal{Y} : i \in \mathbb{N}\} \qquad (*)$$

$$\ker(\mathcal{K})^\perp = \overline{\operatorname{span}}\{u_i \in \mathcal{X} : i \in \mathbb{N}\}$$

and $\mathcal{K} u_i = \sigma_i v_i, \quad \mathcal{K}^* v_i = \sigma_i u_i$

$$\mathcal{K} x = \sum_{i=1}^{\infty} \sigma_i \langle x, u_i \rangle_{\mathcal{X}} v_i \quad \text{for all } x \in \mathcal{X}$$

$$\mathcal{K}^* y = \sum_{i=1}^{\infty} \langle y, v_i \rangle_{\mathcal{Y}} u_i \quad \text{for all } y \in \mathcal{Y}$$

we call triplets $\{(\sigma_i, u_i, v_i)\} \subset]0,\infty[\times \mathcal{X} \times \mathcal{Y}$ a <u>singular system</u> of the operator $K \in \mathcal{L}(\mathcal{X},\mathcal{Y})$

(iii) the equation $Kx = y$ has a solution if and only if

$$y = \sum_{i=1}^{\infty} \langle y, v_i \rangle_y v_i, \qquad \sum_{i=1}^{\infty} |\langle y, v_i\rangle_y|^2/\sigma_i^2 < \infty \quad (*)$$

Picard condition

projection on $\overline{\text{range}(K)}$; see (*)

solution is given by

can be chosen arbitrarily

$$x = x_0 + \sum_{i=1}^{\infty} \langle y, v_i\rangle_y u_i/\sigma_i, \qquad x_0 \in \ker(K)$$

with $x_0 = 0$ we obtain Moore – Penrose inverse

$$x^+ = K^+ y := \sum_{i=1}^{\infty} \langle y, v_i \rangle_y u_i/\sigma_i$$

let $P := \text{proj}_{\overline{\text{range}(K)}}$ denote orthogonal projection on closure of range(K); we have

$$P : \mathcal{Y} \to \overline{\text{range}(K)}, \quad y \mapsto \sum_{i=1}^{\infty} \langle y, v_i\rangle_y v_i$$

it follows that for any $x \in \mathcal{X}$, we have

$$\|Kx - y\|_{\mathcal{Y}}^2 = \|Kx - Py\|_{\mathcal{Y}}^2 + \|(id - P)y\|_{\mathcal{Y}}^2 \geq \|(id-P)y\|_{\mathcal{Y}}^2$$

if y has nonzero components in subspace orthogonal to range(K), $Kx = y$ cannot be satisfied exactly

best we can do is to solve projected equation

$$Kx = PKx = Py \qquad (*)$$

finite dimensional orthogonal projection

$$P_k : \mathcal{Y} \to \text{span}\{v_1, v_2, \dots, v_k\}, \qquad y \mapsto \sum_{i=1}^{k} \langle y, v_i\rangle_y v_i$$

we have that $P_k y \in \text{range}(K)$, and $P_k \to P$ in $\mathcal{Y}$ as $k \to \infty$

we replace (*) by

$$Kx = P_k y, \quad k \in \mathbb{N} \qquad (*)$$

taking on both sides the inner product with v_i we have

$$\langle Kx, v_i\rangle_Y = \langle P_k y, v_i\rangle_Y$$

left hand side: $\langle Kx, v_i\rangle_Y = \langle x, K^* v_i\rangle_X$ and with $K^* v_i = \sigma_i u_i$

we have $\langle x, K^* v_i\rangle_X = \langle x, \sigma_i u_i\rangle_X = \sigma_i \langle x, u_i\rangle_X$

right hand side:

$$\langle P_k y, v_i\rangle_Y = \langle \sum_{j=1}^{k} \langle y, v_j\rangle_Y v_j, v_i\rangle_Y = \sum_{j=1}^{k} \langle y, v_j\rangle_Y \langle v_j, v_i\rangle_Y$$

with $\langle v_j, v_i\rangle_Y = 1$ if $i=j$ and $\langle v_j, v_i\rangle_Y = 0$ otherwise, we have

$$\sum_{j=1}^{k} \langle y, v_j\rangle_Y \langle v_j, v_i\rangle_Y = \langle y, v_i\rangle_Y$$

therefore

$$\sigma_i \langle x, u_i\rangle_X = \begin{cases} \langle y, v_i\rangle_Y, & 1 \le i \le k \\ 0, & i > k \end{cases}$$

solution to (*) is given by (with $x_0 \in \ker(K)$, $x_0 = 0$)

$$x_{TSVD} = \sum_{i=1}^{k} \langle y, v_i\rangle_Y u_i / \sigma_i$$

observe that since for increasing k

$$\|Kx_k - Py\|_Y^2 = \|(P_k - P)y\|_Y^2 \to 0$$

can be made arbitrarily small

[def]

let $K: X \to Y$ be a compact operator with singular value system $\{(\sigma_i, u_i, v_i)\}$; the truncated SVD approximation of the problem $Kx = y$ is: find $x \in X$ such that $Kx = P_k y$, $x \perp \ker(K)$

[thm]

the problem given in definition above has a unique solution

$$x_{TSVD} = \sum_{i=1}^{k} \langle y, v_i\rangle_Y u_i / \sigma_i$$

construction of regularization methods

for $y \in \mathrm{dom}(K^\dagger) = \mathrm{range}(K) \oplus (\mathrm{range}(K))^\perp$ we replace

$$K^\dagger y = \sum_{i=1}^{\infty} \langle y, v_i\rangle_Y u_i / \sigma_i$$

by $R_\alpha : \mathcal{Y} \to \mathcal{X}$

$$R_\alpha y := \sum_{i=1}^{\infty} g_\alpha(\sigma_i) \langle y, v_i \rangle_{\mathcal{Y}} u_i$$

"spectral regularization"

where $g_\alpha : \mathbb{R}_{++} \to \mathbb{R}_+$ is a bounded function

that is $g_\alpha(\sigma) \leq c_\alpha$ for all $\sigma \in \mathbb{R}_{++}$ and fixed $\alpha > 0$

with $\lim_{\alpha \to 0} g_\alpha(t) = 1/t$, i.e., $R_\alpha y \to K^+ y$ for $\alpha \to 0$

with $g_\alpha(\sigma) \leq c_\alpha$ we have

$$\|R_\alpha y\|_{\mathcal{X}}^2 = \sum_{i=1}^{\infty} (g_\alpha(\sigma_i))^2 |\langle y, v_i \rangle_{\mathcal{Y}}|^2 \leq c_\alpha^2 \sum_{i=1}^{\infty} |\langle y, v_i \rangle_{\mathcal{Y}}|^2 \leq c_\alpha^2 \|y\|_{\mathcal{Y}}^2$$

[def]

let $K \in \mathcal{L}(\mathcal{X}, \mathcal{Y})$; a family $\{R_\alpha\}_{\alpha > 0}$ of linear operators is called regularization of K^+ if

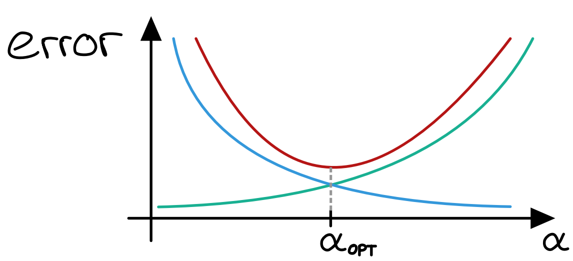
- $R_\alpha \in \mathcal{L}(\mathcal{Y}, \mathcal{X})$ for $\alpha > 0$
- $R_\alpha y \to K^+ y = x^+$ for $\alpha \to 0$ for all $y \in \mathrm{dom}(K^+)$

$$\exists\ y \in \mathrm{dom}(K^+) \text{ such that } \|y - y^\delta\|_{\mathcal{Y}} \leq \delta$$

let $x_\alpha^\delta := R_\alpha y^\delta$, $x^+ := K^+ y$; we have

$$\begin{aligned}\|x_\alpha^\delta - x^+\|_{\mathcal{X}} = \|R_\alpha y^\delta - K^+ y\|_{\mathcal{X}} &= \|R_\alpha y^\delta - R_\alpha y + R_\alpha y - K^+ y\|_{\mathcal{X}} \\ &\leq \|R_\alpha y^\delta - R_\alpha y\|_{\mathcal{X}} + \|R_\alpha y - K^+ y\|_{\mathcal{X}} \\ &\leq \|R_\alpha\|_{\mathcal{L}(\mathcal{Y},\mathcal{X})} \|y - y^\delta\|_{\mathcal{Y}} + \|R_\alpha y - K^+ y\|_{\mathcal{X}} \\ &\leq \underbrace{\delta \|R_\alpha\|_{\mathcal{L}(\mathcal{Y},\mathcal{X})}}_{\text{data error}} + \underbrace{\|R_\alpha y - K^+ y\|_{\mathcal{X}}}_{\text{approximation error}}\end{aligned}$$

- data error does not stay bounded for $\alpha \to 0$
- approximation error vanishes for $\alpha \to 0$



total error

approximation error

data error

$\alpha = \alpha(\delta)$ needs to be chosen so that
(i) approximation error becomes as small as possible
(ii) data error remains controlled

truncated SVD revisited

idea is to discard singular values below certain threshold $\alpha > 0$

$$g_\alpha(\sigma) = \begin{cases} \sigma^{-1} & \text{for } \sigma \geq \alpha \\ 0 & \text{otherwise} \end{cases}$$

note that for all $\sigma > 0$ we obtain

$$\lim_{\alpha \to 0} g_\alpha(\sigma) = \sigma^{-1}$$

regularization scheme is formally given by

$$R_\alpha y := \sum_{\sigma_i \geq \alpha} \sigma_i^{-1} \langle y, v_i \rangle_{\mathcal{Y}} u_i \qquad \text{for all } y \in \mathcal{Y}$$

we observe $g_\alpha(\sigma) \leq \alpha^{-1} =: C_\alpha$ so that $\|R_\alpha\|_{\mathcal{L}(\mathcal{Y},\mathcal{X})} \leq C_\alpha$

Tikhonov regularization

idea is to shift singular values of K^*K by constant factor

$$g_\alpha(\sigma) = \frac{\sigma}{\sigma^2 + \alpha}$$

note that for all $\sigma > 0$ we obtain

$$\lim_{\alpha \to 0} g_\alpha(\sigma) = \sigma^{-1}$$

regularization scheme is formally given by

$$R_\alpha y := \sum_{i=1}^{\infty} \frac{\sigma_i}{\sigma_i^2 + \alpha} \langle y, v_i \rangle_{\mathcal{Y}} u_i \qquad \text{for all } y \in \mathcal{Y} \qquad (*)$$

since $\sigma_i^2 + \alpha \geq 2\sqrt{\alpha}\,\sigma$, we estimate

$$g_\alpha(\sigma) \leq (2\sqrt{\alpha})^{-1} =: C_\alpha$$

so that $\|R_\alpha\|_{\mathcal{L}(\mathcal{Y},\mathcal{X})} \leq C_\alpha$

we do not need to compute SVD to apply Tikhonov regularization

SVD representation of $(K^*K+\alpha\, id)x_\alpha$, $x_\alpha := R_\alpha y$:

$$(K^*K+\alpha\, id)x_\alpha = (K^*K+\alpha\, id)R_\alpha y$$

with (*) we have

$$=(K^*K+\alpha\, id)\sum_{i=1}^{\infty}\frac{\sigma_i}{\sigma_i^2+\alpha}\langle y,v_i\rangle_y u_i$$

$$=\sum_{i=1}^{\infty}\frac{\sigma_i}{\sigma_i^2+\alpha}\langle y,v_i\rangle_y K^*K u_i+\sum_{i=1}^{\infty}\frac{\alpha\sigma_i}{\sigma_i^2+\alpha}\langle y,v_i\rangle_y u_i$$

with $Ku_i=\sigma_i v_i$, $K^*v_i=\sigma_i u_i$

we obtain $K^*Ku_i=\sigma_i K^*v_i=\sigma_i^2 u_i$ and by that

$$=\sum_{i=1}^{\infty}\frac{\sigma_i}{\sigma_i^2+\alpha}\langle y,v_i\rangle_y \sigma_i^2 u_i+\sum_{i=1}^{\infty}\frac{\alpha\sigma_i}{\sigma_i^2+\alpha}\langle y,v_i\rangle_y u_i$$

$$=\sum_{i=1}^{\infty}\frac{\sigma_i(\sigma_i^2+\alpha)}{\sigma_i^2+\alpha}\langle y,v_i\rangle_y u_i = K^*y$$

let $K:\mathcal{X}\to\mathcal{Y}$ be a compact operator with singular value system $\{(\sigma_i,u_i,v_i)\}$; then the Tikhonov-regularized solution exists, is unique, and given by **[thm]**

$$x_\alpha=(K^*K+\alpha\, id)^{-1}K^*y=\sum_{i=1}^{\infty}\frac{\sigma_i}{\sigma_i^2+\alpha}\langle y,v_i\rangle_y u_i$$

variational formulation:

for $y\in\mathcal{Y}$ the Tikhonov-regularized solution **[thm]**

$$x_\alpha=R_\alpha y:=\sum_{i=1}^{\infty}\frac{\sigma_i}{\sigma_i^2+\alpha}\langle y,v_i\rangle_y u_i$$

is uniquely determined as the global minimizer of the Tikhonov-functional

$$f_\alpha(x):=\frac{1}{2}\|Kx-y\|_{\mathcal{Y}}^2+\frac{\alpha}{2}\|x\|_{\mathcal{X}}^2$$

proof:

global minimizer $\tilde{x} \in X$ of f_α is characterized by

$$f_\alpha(\tilde{x}) \leq f_\alpha(x) \quad \text{for all } x \in X$$

($\Longrightarrow$) let x_α be Tikhonov-regularized solution

we have

$$\begin{aligned} f_\alpha(x) - f_\alpha(x_\alpha) &= \tfrac{1}{2}\langle Kx, Kx\rangle_Y - \langle Kx, y\rangle_Y + \cancel{\tfrac{1}{2}\langle y, y\rangle_Y} + \tfrac{\alpha}{2}\langle x, x\rangle_X \\ &\quad - \tfrac{1}{2}\langle Kx_\alpha, Kx_\alpha\rangle_Y + \langle Kx_\alpha, y\rangle_Y - \cancel{\tfrac{1}{2}\langle y, y\rangle_Y} - \tfrac{\alpha}{2}\langle x_\alpha, x_\alpha\rangle_X \\ &\quad + \underbrace{\langle (K^*K + \alpha\, id)x_\alpha - K^*y, x_\alpha - x\rangle_X}_{=0} \\ &= \tfrac{1}{2}\|Kx - Kx_\alpha\|_Y^2 + \tfrac{\alpha}{2}\|x - x_\alpha\|_X^2 \geq 0 \end{aligned}$$

implies x_α is a global minimizer of f_α

($\Longleftarrow$) let $\tilde{x} \in X$ be a global minimizer with

$$f_\alpha(\tilde{x}) \leq f_\alpha(x) \quad \text{for all } x \in X$$

it follows for $x = \tilde{x} + \tau z$ with arbitrary $\tau > 0$ and fixed $z \in X$ that

$$\begin{aligned} 0 \leq f_\alpha(x) - f_\alpha(\tilde{x}) &= \tfrac{\tau^2}{2}\|Kz\|_Y^2 + \tfrac{\alpha\tau^2}{2}\|z\|_X^2 \\ &\quad + \tau\langle (K^*K + \alpha\, id)\tilde{x} - K^*y, z\rangle_X \end{aligned}$$

dividing by τ and taking the limit $\tau \to 0$ yields

$$\begin{aligned} &\lim_{\tau\to 0} \tfrac{1}{\tau}\left(\tfrac{\tau^2}{2}\|Kz\|_Y^2 + \tfrac{\alpha\tau^2}{2}\|z\|_X^2 + \tau\langle (K^*K + \alpha\, id)\tilde{x} - K^*y, z\rangle_X\right) \\ &= \langle (K^*K + \alpha\, id)\tilde{x} - K^*y, z\rangle_X \end{aligned}$$

this shows that $\langle (K^*K + \alpha\, id)\tilde{x} - K^*y, z\rangle_X \geq 0$ for all $z \in X$

$\Rightarrow$ $\tilde{x} = x_\alpha$ and x_α is a unique minimizer

generalization

(i) $$\tilde{f}_\alpha(x) = \tfrac{1}{2}\|Kx - y\|_Y^2 + \alpha\, reg(x)$$

nonlinear

we cannot expect that

$$x_\alpha = \arg\min_{x\in X} \tilde{f}_\alpha(x) \to x^\dagger \quad \text{for } \alpha \to 0$$

(ii) $\tilde{f}_\alpha(x) = \frac{1}{2}\|K(x) - y\|_Y^2 + \alpha \operatorname{reg}(x)$

nonlinear

parameter-choice strategies

we will consider 3 rules for choice of α and if they lead to convergent regularization method

[def]

function $\alpha: \mathbb{R}_{++} \times Y \to \mathbb{R}_{++}$, $(\delta, y^\delta) \mapsto \alpha(\delta, y^\delta)$ is called parameter choice rule

(i) a-priori parameter choice rules, if they depend on δ only

(ii) a-posteriori parameter choice rules, if they depend on δ and y^δ

(iii) heuristic parameter choice rules, if they depend on y^δ only

[def]

let $\{R_\alpha\}_{\alpha>0}$ be a regularization of $K^\dagger$ and $\alpha: \mathbb{R}_{++} \times Y \to \mathbb{R}_{++}$ be a parameter choice rule; we call pair (R_α, α) convergent regularization, if for all $y \in \operatorname{dom}(K^\dagger)$, there exists a parameter choice rule such that

"limit superior"

$$\limsup_{\delta \to 0} \{\|R_\alpha y^\delta - K^\dagger y\|_X : y^\delta \in Y, \|y - y^\delta\|_Y \leq \delta\} = 0$$

$$\limsup_{\delta \to 0} \{\alpha(\delta, y^\delta) : y^\delta \in Y, \|y - y^\delta\|_Y \leq \delta\} = 0$$

are guaranteed

a priori parameter choice rules

[thm]

let $\{R_\alpha\}_{\alpha>0}$ be a regularization of $K^\dagger$ for $K \in \mathcal{L}(X, Y)$ and $\alpha: \mathbb{R}_{++} \to \mathbb{R}_{++}$; there exists an a-priori parameter choice $\alpha = \alpha(\delta)$ such that (R_α, α) is convergent proof: let $y \in \mathrm{dom}(K^\dagger)$; we can find a monotone increasing function $\beta: \mathbb{R}_{++} \to \mathbb{R}_{++}$, $\lim_{\varepsilon\to 0} \beta(\varepsilon) = 0$, such that $\|R_{\beta(\varepsilon)} y - K^\dagger y\|_X \le \frac{\varepsilon}{2}$ due to pointwise convergence of $R_\alpha \to K^\dagger$

since $R_{\beta(\varepsilon)}$ is continous for fixed ε, $\exists \rho(\varepsilon) > 0$ with $\|R_{\beta(\varepsilon)} y - R_{\beta(\varepsilon)} \tilde{y}\|_X \le \frac{\varepsilon}{2}$ for all $\tilde{y} \in \mathcal{Y}$ where $\|y - \tilde{y}\|_Y \le \rho(\varepsilon)$

WLOG, let $\rho: \mathbb{R}_{++} \to \mathbb{R}_{++}$ be continous, strictly monotone increasing function with $\lim_{\varepsilon\to 0} \rho(\varepsilon) = 0$

by the inverse function theorem there exists a continuous strictly monotone inverse $\rho^{-1}: \mathbb{R}_{++} \to \mathbb{R}_{++}$ on range(ρ) with $\lim_{\delta\to 0} \rho^{-1}(\delta) = 0$

we continously extend ρ^{-1} on $\mathbb{R}_{++}$ and define the a priori strategy as $\alpha: \mathbb{R}_{++} \to \mathbb{R}_{++}$, $\delta \to \beta(\rho^{-1}(\delta))$, $\lim_{\delta \to 0} \alpha(\delta) = 0$

moreover, $\exists\, \delta := \rho(\varepsilon)$ for all $\varepsilon > 0$, with $\alpha(\delta) = \beta(\varepsilon)$, such that

$$\begin{aligned}\|R_\alpha y^\delta - K^\dagger y\|_X &= \|R_{\beta(\varepsilon)} y^\delta - R_{\beta(\varepsilon)} y + R_{\beta(\varepsilon)} y - K^\dagger y\|_X \\ &\le \underbrace{\|R_{\beta(\varepsilon)} y^\delta - R_{\beta(\varepsilon)} y\|_X}_{\le \frac{\varepsilon}{2}} + \underbrace{\|R_{\beta(\varepsilon)} y - K^\dagger y\|_X}_{\le \frac{\varepsilon}{2}} \le \varepsilon\end{aligned}$$

follows for all $\|y - y^\delta\|_Y \le \delta$

$\Rightarrow (R_\alpha, \alpha)$ is convergent

[thm]

let $\{R_\alpha\}_{\alpha>0}$ be a linear regularization and $\alpha : \mathbb{R}_{++} \to \mathbb{R}_{++}$ an a-priori parameter choice rule; (R_α, α) is convergent iff

(i) $\lim_{\delta \to 0} \alpha(\delta) = 0$

(ii) $\lim_{\delta \to 0} \delta \|R_\alpha\|_{\mathcal{L}(\mathcal{Y},\mathcal{X})} = 0$

proof: ($\Leftarrow$) let (i) and (ii) hold; since (see above)

$$\|R_\alpha y^\delta - K^+ y\|_\mathcal{X} \le \delta \|R_\alpha\|_{\mathcal{L}(\mathcal{Y},\mathcal{X})} + \|R_\alpha y - K^+ y\|_\mathcal{X}$$

we observe that

$$\|R_\alpha y^\delta - K^+ y\|_\mathcal{X} \to 0 \text{ as } \delta \to 0$$

for any $y^\delta \in \mathcal{Y}$, $y \in \operatorname{dom}(K^+)$ such that

$$\|y - y^\delta\|_\mathcal{Y} \le \delta$$

($\Rightarrow$) exercise ...

typical parameter choice strategies are of form $\alpha(\delta) = \delta^p$

identifying optimal values p requires additional info about x^+

if a priori parameter choice rule $\alpha(\delta)$ defines convergent regularization, then so does $\tilde{\alpha}(\delta) = \alpha(c\delta)$ for any $c > 0$

a-posteriori parameter choice

basic idea: let $y \in \operatorname{dom}(K^+)$, $\|y - y^\delta\|_\mathcal{Y} \le \delta$, $x_\alpha^\delta := R_\alpha y^\delta$

consider residual

$$\|K x_\alpha^\delta - y^\delta\|_\mathcal{Y} = \|K R_\alpha y^\delta - y^\delta\|_\mathcal{Y}$$

let x^+ be minimal norm solution and let

$$\mu := \inf_{x \in \mathcal{X}} \{\|Kx - y\|_\mathcal{Y}\} = \|Kx^+ - y\|_\mathcal{Y}$$

then we observe that x^+ satisfies

$$\|Kx^+ - y^\delta\|_\mathcal{Y} \le \|Kx^+ - y\|_\mathcal{Y} + \|y - y^\delta\|_\mathcal{Y} \le \mu + \delta$$

in some cases this estimate might be sharp

it appears to be useful to select $\alpha(\delta, y^\delta)$ such that

$$\|Kx_\alpha^\delta - y^\delta\|_Y = \|KR_\alpha y^\delta - y^\delta\|_Y \leq \mu + \delta$$

best case scenario for comparison

$$\|Kx^\dagger - y^\delta\|_Y \underset{Kx^\dagger = y}{=} \|y - y^\delta\| = \delta, \text{ i.e., } \mu = 0$$

for given data y^δ it does not make sense to identify $\alpha(\delta, y^\delta)$ with a residuum beyond data error δ, i.e.,

$$\|Kx_{\alpha(\delta,y^\delta)} - y^\delta\|_Y = \|KR_{\alpha(\delta,y^\delta)} y^\delta - y^\delta\|_Y \approx \delta$$

leads to Morozov's discrepancy principle:

[def]

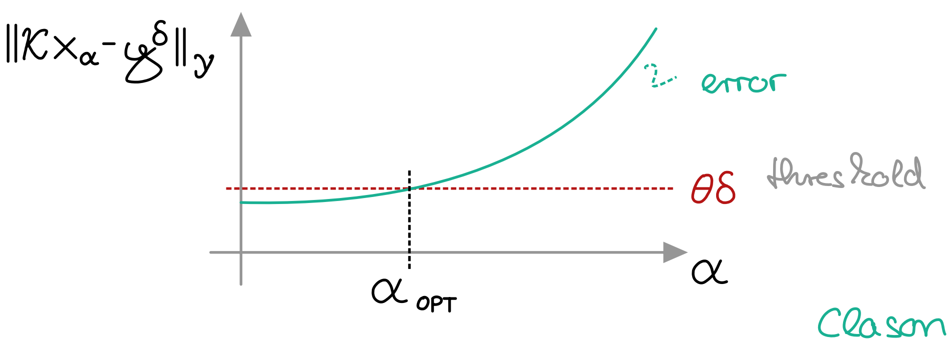
let $\delta > 0$, $y^\delta \in Y$, $\theta > 1$; we select $\alpha(\delta, y^\delta)$ such that

$$\alpha(\delta, y^\delta) = \sup\{\alpha > 0 \; \|Kx_\alpha - y^\delta\|_Y = \|KR_\alpha y^\delta - y^\delta\|_Y \leq \theta\delta\}$$

for given $y^\delta \in Y$, $\delta > 0$, and fixed constant $\theta > 1$; then

$$x_{\alpha(\delta,y^\delta)} = R_{\alpha(\delta,y^\delta)} y^\delta$$

is said to satisfy Morozov's discrepancy principle



heuristic parameter choice rules

"Bakushinskii veto" [thm]

let $\{R_\alpha\}_{\alpha>0}$ be a regularization of $K^\dagger$ for $K \in \mathcal{L}(X, Y)$ and $\alpha : Y \to \mathbb{R}_{++}$, $\alpha = \alpha(y^\delta)$ be a parameter choice rule such that (R_α, α) is convergent; then $K^\dagger$ is continuous proof: suppose such a rule $\alpha = \alpha(y)$ exists; then we can define a map

$$\mathcal{R}_\alpha : \mathcal{Y} \to \mathcal{X}, \quad y \mapsto \mathcal{R}_{\alpha(y)} y$$

let $y \in \mathrm{dom}(\mathcal{K}^+)$ be arbitrary; since we assume that $(\mathcal{R}_\alpha, \alpha)$ is convergent we have

(*) $\lim\sup_{\delta \to 0} \{ \|\mathcal{R}_\alpha y^\delta - \mathcal{K}^+ y\|_\mathcal{X} : y^\delta \in \mathcal{Y}, \|y^\delta - y\|_\mathcal{Y} \leq \delta \} = 0$

from this, we have with $y^\delta = y$ and $\delta = 0$ that $\mathcal{R}_\alpha y = \mathcal{K}^+ y$ for all $y \in \mathrm{dom}(\mathcal{K}^+)$

for an arbitrary sequence $\{y_i\}_{i \in \mathbb{N}} \subset \mathrm{dom}(\mathcal{K}^+)$, with $y_i \to y$, (*) together with $y_i = y^\delta$ and $\delta := \|y^\delta - y\|_\mathcal{Y}$ that $\mathcal{K}^+ y_i = \mathcal{R}_\alpha y_i = \mathcal{R}_{\alpha(y_i)} y_i \to \mathcal{R}_{\alpha(y)} y \to \mathcal{K}^+ y$

$\Rightarrow \mathcal{K}^+$ is continuous

- quasi optimality principle: for first n elements of zero sequence $\{\alpha_i\}_{i \in \mathbb{N}}$, we choose $\alpha(y^\delta) = \alpha_{i_{OPT}}$, with

$$i_{OPT} = \arg\min_{1 \leq i < n} \|x_{\alpha_{i+1}} - x_{\alpha_i}\|_\mathcal{X}$$

- Hanke – Raus rule:

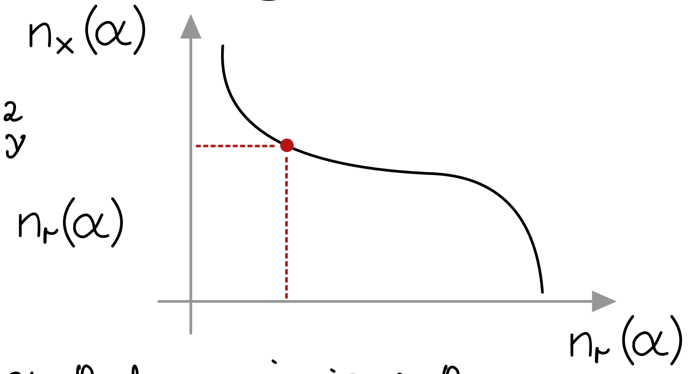
$$\alpha(y^\delta) = \arg\min_{\alpha \in \mathbb{R}_{++}} (\sqrt{\alpha})^{-1} \|\mathcal{K} x_\alpha - y^\delta\|_\mathcal{Y}$$

- L-curve:

$$\alpha(y^\delta) = \arg\min_{\alpha \in \mathbb{R}_{++}} \|x_\alpha\|_\mathcal{X} \|\mathcal{K} x_\alpha - y^\delta\|_\mathcal{Y}$$

let $n_x(\alpha) := \log \|x_\alpha\|_\mathcal{X}^2$

$n_r(\alpha) := \log \|\mathcal{K} x_\alpha - y^\delta\|_\mathcal{Y}^2$



we assume that $n_x(\alpha)$ and $n_r(\alpha)$ vary smoothly with α

one selects the value of α that maximizes the curvature function

$$\kappa(\alpha) := \frac{n_r''(\alpha) n_x'(\alpha) - n_r'(\alpha) n_x''(\alpha)}{(n_r'(\alpha)^2 + n_x'(\alpha)^2)^{3/2}}$$