Regularization

- ill-conditioned problems
- noise on the data

⇒ without regularization
computed solution is useless

⇒ under certain conditions,
regularization can be used

Example: integral equation

$$\int_a^b f(y)\,K(x,y)\,dy = g(x)$$

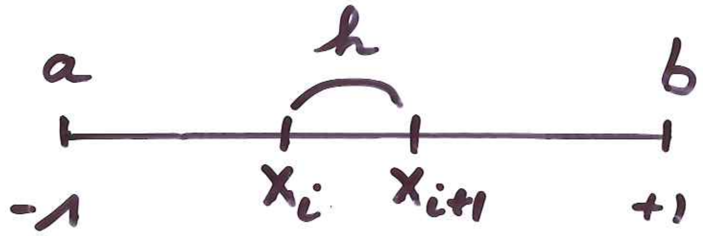
! $g(x)$, $K(x,y)$? $f(y)$

Fredholm integral equation of the first kind

e.g. $$\int_{-1}^{+1} f(y)\, e^{-\alpha(x-y)^2}\,dy = g(x)$$

the function $f(y)$ is somehow smoothed into the function $g(x)$

discretisation:



m intervals: $h = \dfrac{b-a}{m}$

$$x_i = a + \left(i - \frac{1}{2}\right) \cdot h \quad , \quad i = 1(1)m$$

$$\Rightarrow \quad g(x_i) = \frac{1}{h} \sum_{j=1}^{m} K(x_i, y_j) f(y_j)$$

(mid-point rule)

$\Rightarrow$ linear system of equations:

$$A x = b$$

! $A = [K(x_i, y_j)]$? $x = [f(y_j)]$

$b = [g(x_i)]$

More general formulation:

$$A : m \times n \quad : m \geq n$$

$$b : m \times 1$$

$$x : n \times 1$$

least squares criterion:

$$\| A x - b \|_2 \swarrow$$

singular value decomposition of A

$$A = [U_1 \;\; U_2] \begin{bmatrix} \Sigma \\ 0 \end{bmatrix} V^H$$

$$\Rightarrow \left\| [U_1 \;\; U_2] \begin{bmatrix} \Sigma \\ 0 \end{bmatrix} V^H x - b \right\|_2 \swarrow$$

$$\Rightarrow \left\| \begin{bmatrix} \Sigma \\ 0 \end{bmatrix} \underbrace{V^H x}_{\substack{\parallel \\ y}} - \begin{bmatrix} U_1^H b \\ U_2^H b \end{bmatrix} \right\|_2 \swarrow$$

$$y_i = \frac{u_i^H b}{\sigma_i} \quad , \; i = 1(1)n$$

$$x = Vy = \sum_{i=1}^{n} \frac{u_i^H b}{\sigma_i} v_i$$

Tikhonov regularization

$$x_\lambda = \underset{x}{\operatorname{argmin}} \; \|Ax - b\|_2^2 + \|\lambda x\|_2^2$$

$$= \underset{x}{\operatorname{argmin}} \left\| \begin{bmatrix} A \\ \lambda I \end{bmatrix} x - \begin{bmatrix} b \\ 0 \end{bmatrix} \right\|_2$$

$$= \underset{x}{\operatorname{argmin}} \left\| \begin{bmatrix} \Sigma \\ 0 \\ \lambda I \end{bmatrix} \underbrace{V^H x}_{y} - \begin{bmatrix} U_1^H b \\ U_2^H b \\ 0 \end{bmatrix} \right\|_2$$

$$\Rightarrow f_i(y_i) = (\sigma_i y_i - u_i^H b)^2 + \lambda^2 y_i^2$$

parabola in y_i

$$\frac{d f_i(y_i)}{d y_i} = 0 \quad : \quad \cancel{2}(\sigma_i y_i - u_i^H b)\sigma_i + \cancel{2}\lambda^2 y_i = 0$$

$$\Rightarrow y_i = \frac{\sigma_i^2}{\sigma_i^2 + \lambda^2} \left(\frac{u_i^H b}{\sigma_i} \right)$$

filter factors

$$x_\lambda = \sum_{i=1}^{m} \underbrace{\frac{\sigma_i^2}{\sigma_i^2+\lambda^2}}_{f_i} \cdot \left(\frac{u_i^H b}{\sigma_i}\right) \cdot v_i$$

$$b - A x_\lambda = b_0 + \sum_{i=1}^{m} (u_i^H b)\, u_i$$

$$+ A \cdot \sum_{i=1}^{m} \frac{\sigma_i^2}{\sigma_i^2+\lambda^2} \cdot \left(\frac{u_i^H b}{\sigma_i}\right) v_i$$

$$= b_0 + \sum_{i=1}^{m} (1 - f_i)\,(u_i^H b)\, u_i$$

SVD - analysis

$$b = \bar{b} + e \qquad \bar{b} = A\bar{x}$$

errors in the data

exact RHS

exact solution

theory: $\bar{x}_\lambda = (A^TA + \lambda^2 I)^{-1} A^T \bar{b}$

$+$

$x_\lambda^e = (A^TA + \lambda^2 I)^{-1} A^T e$

---

x_λ

discrete Picard condition:

exact SVD coefficients

$|u_i^H \bar{b}|$ decay faster than σ_i

$\Rightarrow |v_i^H \bar{x}| = |u_i^T \bar{b} / \sigma_i|$ decays

$\Rightarrow \bar{x}$ has not a large norm

e : error : white noise

$$E\left((u_i^T e)^2\right) = \varepsilon^2 , \quad i = 1(1)m$$

→ does not satisfy the discrete Picard condition

$$\|x_\lambda^e\|_2^2 \simeq \sum_{i=1}^{m} \left(\frac{\sigma_i \, \varepsilon}{\sigma_i^2 + \lambda^2}\right)^2 \qquad \lambda \approx \sigma_k$$

$$\simeq \sum_{i=1}^{k} \left(\frac{\varepsilon}{\sigma_i}\right)^2 + \sum_{i=k+1}^{m} \left(\frac{\sigma_i \, \varepsilon}{\lambda^2}\right)^2$$

$$\simeq \varepsilon^2 \left[\underbrace{\sum_{i=1}^{k} \sigma_i^{-2}} + \lambda^{-4} \underbrace{\sum_{i=k+1}^{m} \sigma_i^2} \right]$$

dominated by: $\sigma_i^{-2} \approx \lambda^{-2}$ $\qquad$ $\sigma_{k+1}^2 \approx \lambda^2$

Hence, $\|x_\lambda^e\|_2 \approx \varepsilon . \frac{c_\lambda}{\lambda}$ → varies slowly with λ

$$\|A x_\lambda^e - e\|_2^2 \simeq \sum_{i=k+1}^{m} \varepsilon^2 = (m-k)\,\varepsilon^2$$

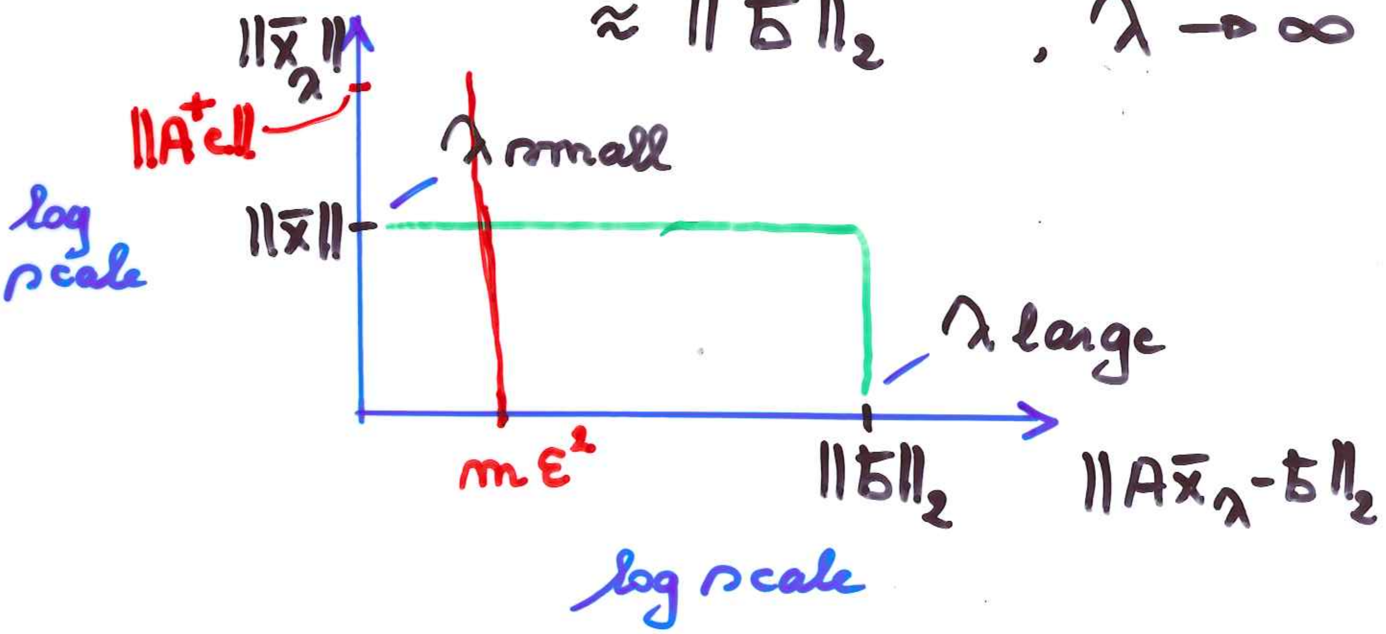
λ not large

$$\|\bar{x}_\lambda\|_2 \approx \|\bar{x}\|_2$$

$$\lambda \to \infty \quad : \quad \|\bar{x}_\lambda\|_2 \to 0$$

$$\|A\bar{x}_\lambda - b\|_2 = \|b_0\|_2 = 0 \quad \text{when } \lambda = 0$$

$$\approx \|b\|_2 \quad , \lambda \to \infty$$

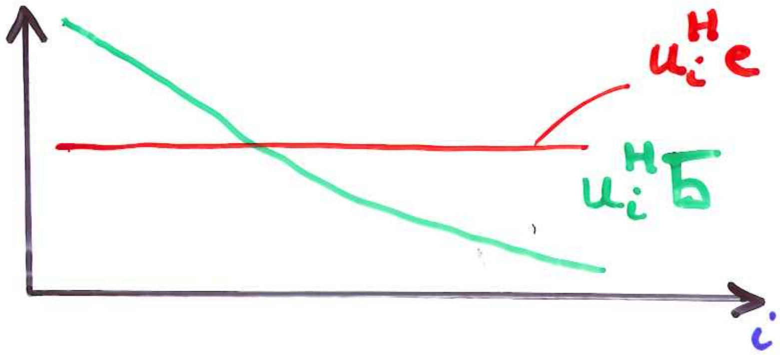


L-curve

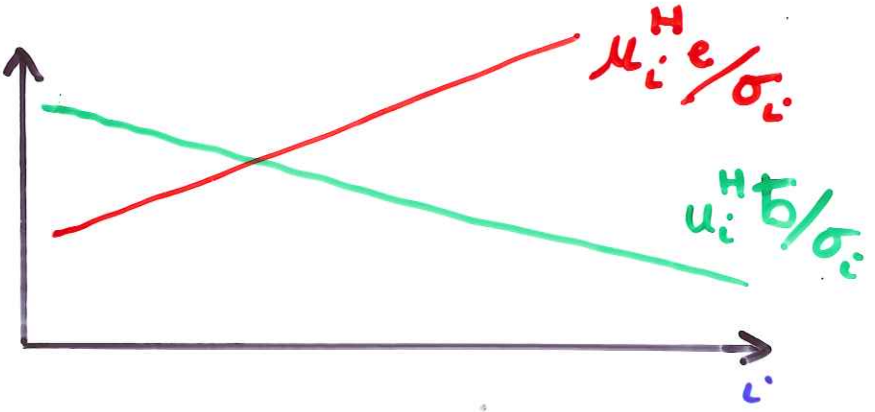
Intuitive way of reasoning

$$b = \bar{b} + e$$

log



log



filter: Tikhonov regularization

$$f_i = \frac{\sigma_i^2}{\sigma_i^2 + \lambda^2} = \frac{1}{1 + \left(\frac{\lambda}{\sigma_i}\right)^2}$$