Inverse Problems

Michaelmas term 2020

Yury Korolev and Jonas Latz

8 October 2020

UNIVERSITY OF CAMBRIDGE

Practicalities

- Lectures take place on Tuesday, Thursday and Saturday, 11am-12pm, online
- Course materials (lecture notes, example sheets, etc.) will be provided at `http://www.damtp.cam.ac.uk/research/cia/inverse-problems-michaelmas-2020`
- Lectures will be recorded;
- Four example sheets and example classes (possibly in person): 28 October, 18 November, 2 December & some time in January
- For further questions email y.korolev@maths.cam.ac.uk or jl2160@cam.ac.uk

Example classes

Example classes will be held (most likely) online on the following dates:

- Wednesday 28 October, 1.30-3.00pm
- Wednesday 18 November, 1.30-3.00pm
- Wednesday 2 December, 1.30-3.00pm
- Wednesday 20 January, 1.30-3.00pm

Return solutions to two questions (specified in advance)

- Example sheets will be made available one week before the deadline.
- Upload your answers on Moodle by 12 noon on Monday.

Office hours will be held online on Wednesdays 1:30-2:30pm. Please send us an email in advance to arrange a meeting.

What do we mean by inverse problems?

- Direct problem: Given an object (cause), determine data (effect).
- Inverse problem: Observing (noisy) data, recover the object.

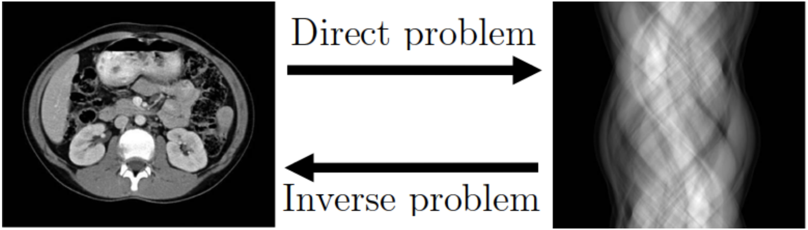
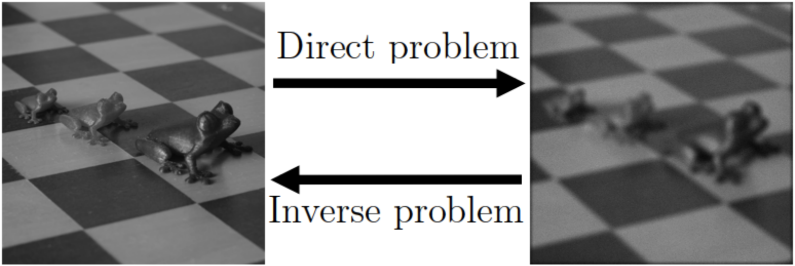


Image processing is a classical example of an inverse problem



Inverse problems are ill-posed

Well-posedness of a problem as defined by Jacques Hadamard (1923):

I) Existence. There should be at least one solution.

II) Uniqueness. There should be at most one solution.

III) Stability. The solution must depend continuously on data.

If any of the above conditions is violated the problem is called ill-posed

Linear inverse problems

We consider the linear inverse problem

$$f = Au$$

- The physical phenomenon that relates the unknown and the measurement is modelled by a linear operator $A : X \to Y$.
- The ideal measurement is given by $f = Au$.

Linear inverse problems

We consider the linear inverse problem

$$f = Au$$

- The physical phenomenon that relates the unknown and the measurement is modelled by a linear operator $A : X \to Y$.
- The ideal measurement is given by $f = Au$.
- We can only observe noisy measurements $f_n = Au + n$

We would like to recover the unknown u from a noisy measurement f_n.

Linear inverse problems

We consider the linear inverse problem

$$f = Au$$

- The physical phenomenon that relates the unknown and the measurement is modelled by a linear operator $A : X \to Y$.
- The ideal measurement is given by $f = Au$.
- We can only observe noisy measurements $f_n = Au + n$

We would like to recover the unknown u from a noisy measurement f_n.

Main difficulty: A^{-1} does not exist or is not continuous.

Examples: matrix inversion

I) $A : \mathbb{R}^d \to \mathcal{R}(A) \subsetneq \mathbb{R}^k$, $k > d$, i.e. the system is overdetermined. Because of the noise $f_n \notin \mathcal{R}(A)$
$\Rightarrow$ There is no solution.

Examples: matrix inversion

I) $A : \mathbb{R}^d \to \mathcal{R}(A) \subsetneq \mathbb{R}^k$, $k > d$, i.e. the system is overdetermined. Because of the noise $f_n \notin \mathcal{R}(A)$
$\Rightarrow$ There is no solution.

II) $A : \mathbb{R}^d \to \mathbb{R}^k$, $k < d$, i.e. the system is underdetermined.
$\Rightarrow$ There are several possible solutions.

Examples: matrix inversion

I) $A : \mathbb{R}^d \to \mathcal{R}(A) \subsetneq \mathbb{R}^k$, $k > d$, i.e. the system is overdetermined. Because of the noise $f_n \notin \mathcal{R}(A)$
$\Rightarrow$ There is no solution.

II) $A : \mathbb{R}^d \to \mathbb{R}^k$, $k < d$, i.e. the system is underdetermined.
$\Rightarrow$ There are several possible solutions.

III) There exists $A^{-1} : \mathbb{R}^d \to \mathbb{R}^d$, but the condition number $\kappa = \lambda_1/\lambda_d$ is very large. Then A is almost singular and $\|A^{-1}n\| \approx \|n\|/\lambda_d$ can be arbitrarily large.
$\Rightarrow$ The naive reconstruction $\widetilde{u} = A^{-1}f_n = u + A^{-1}n$ is dominated by the noise.

Examples: Deblurring (deconvolution)

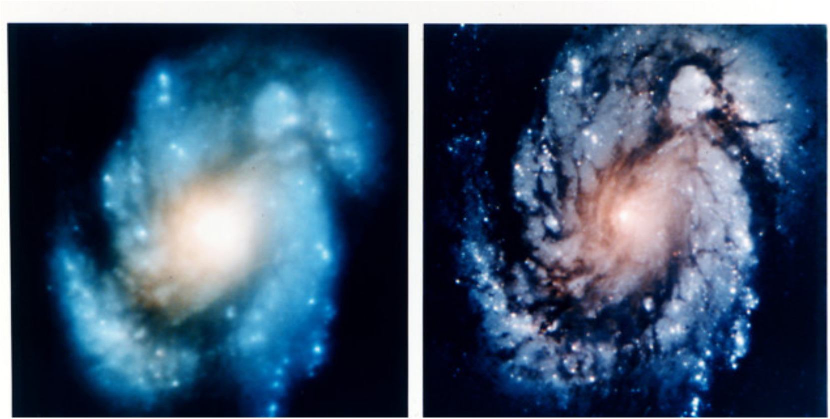


Figure: The Hubble space telescope had a flaw in its mirror which resulted in the images being blurred.

$$f(x) = (Au)(x) = \int_{\mathbb{R}^2} a(x-y)u(y)dy$$

Signal deblurring for noiseless data

The noiseless data $f(t) = \int_{-\infty}^{\infty} a(t-s)u(s)ds$ has Fourier transform

$$\hat{f}(\xi) = \int_{-\infty}^{\infty} e^{-i\xi t} f(t) dt.$$

Signal deblurring for noiseless data

The noiseless data $f(t) = \int_{-\infty}^{\infty} a(t-s)u(s)ds$ has Fourier transform

$$\hat{f}(\xi) = \int_{-\infty}^{\infty} e^{-i\xi t} f(t)dt.$$

The Convolution Theorem implies

$$\hat{f}(\xi) = \hat{a}(\xi)\hat{u}(\xi),$$

Signal deblurring for noiseless data

The noiseless data $f(t) = \int_{-\infty}^{\infty} a(t-s)u(s)ds$ has Fourier transform

$$\hat{f}(\xi) = \int_{-\infty}^{\infty} e^{-i\xi t} f(t)dt.$$

The Convolution Theorem implies

$$\hat{f}(\xi) = \hat{a}(\xi)\hat{u}(\xi),$$

and hence by inverse Fourier transform

$$u(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{it\xi} \frac{\hat{f}(\xi)}{\hat{a}(\xi)} d\xi.$$

Signal deblurring for noisy data

We can only observe noisy data and get

$$\hat{f}_n(\xi) = \hat{a}(\xi)\hat{u}(\xi) + \hat{n}(\xi).$$

Signal deblurring for noisy data

We can only observe noisy data and get

$$\hat{f}_n(\xi) = \hat{a}(\xi)\hat{u}(\xi) + \hat{n}(\xi).$$

We assume that a is a Gaussian kernel

$$a(t) = \frac{1}{\sqrt{2\pi\alpha^2}} \exp\Big(-\frac{1}{2\alpha^2}t^2\Big).$$

Signal deblurring for noisy data

We can only observe noisy data and get

$$\hat{f}_n(\xi) = \hat{a}(\xi)\hat{u}(\xi) + \hat{n}(\xi).$$

We assume that a is a Gaussian kernel

$$a(t) = \frac{1}{\sqrt{2\pi\alpha^2}} \exp\Big(-\frac{1}{2\alpha^2}t^2\Big).$$

Then the estimate $\widetilde{u}$ given by the Convolution Theorem is

$$\widetilde{u}(t) = u(t) + \frac{1}{2\pi}\int_{-\infty}^{\infty} \hat{n}(\xi) \exp\Big(it\xi + \frac{\alpha^2}{2}\xi^2\Big) d\xi,$$

which may not be even well defined since the Fourier transform of the noise will not decay fast enough.

Examples: Heat distribution in an insulated rod

Consider the problem

$$\begin{aligned} v_t - v_{xx} &= 0 \quad \text{in } (0,\pi) \times \mathbb{R}_+ \\ v(0,\cdot) = v(\pi,\cdot) &= 0 \quad \text{on } \mathbb{R}_+ \\ v(\cdot, T) &= f \quad \text{in } (0,\pi) \\ v(\cdot, 0) &= u \quad \text{in } (0,\pi) \end{aligned}$$

Examples: Heat distribution in an insulated rod

Consider the problem

$$\begin{aligned} v_t - v_{xx} &= 0 \quad \text{in } (0,\pi) \times \mathbb{R}_+ \\ v(0,\cdot) = v(\pi,\cdot) &= 0 \quad \text{on } \mathbb{R}_+ \\ v(\cdot, T) &= f \quad \text{in } (0,\pi) \\ v(\cdot, 0) &= u \quad \text{in } (0,\pi) \end{aligned}$$

Forward problem: Determine the final distribution $v(\cdot, T) \in L^2(0,\pi)$, $T > 0$, when the initial distribution $v(\cdot, 0) \in L^2(0,\pi)$ is given.

Examples: Heat distribution in an insulated rod

Consider the problem

$$\begin{aligned} v_t - v_{xx} &= 0 \quad \text{in } (0,\pi) \times \mathbb{R}_+ \\ v(0,\cdot) = v(\pi,\cdot) &= 0 \quad \text{on } \mathbb{R}_+ \\ v(\cdot,T) &= f \quad \text{in } (0,\pi) \\ v(\cdot,0) &= u \quad \text{in } (0,\pi) \end{aligned}$$

Forward problem: Determine the final distribution $v(\cdot,T) \in L^2(0,\pi)$, $T > 0$, when the initial distribution $v(\cdot,0) \in L^2(0,\pi)$ is given.

Inverse problem: Determine the initial distribution $v(\cdot,0) \in L^2(0,\pi)$ from observed (noisy) final distribution $v(\cdot,T) \in L^2(0,\pi)$.

Forward problem

The solution to the forward problem can be given explicitly:

$$v(x, T) = \sum_{j=1}^{\infty} \widehat{u}_j e^{-j^2 T} sin(jx),$$

where $\{\widehat{u}_j\}_{j=0}^{\infty}$ are the Fourier (sine) coefficients of the initial heat distribution u.

Forward problem

The solution to the forward problem can be given explicitly:

$$v(x,T) = \sum_{j=1}^{\infty} \widehat{u}_j e^{-j^2 T} sin(jx),$$

where $\{\widehat{u}_j\}_{j=0}^{\infty}$ are the Fourier (sine) coefficients of the initial heat distribution u.

The solution operator

$$A_T : u \mapsto v(\cdot, T) = f, \quad L^2(0,\pi) \to L^2(0,\pi)$$

satisfies the following conditions:

- A_T is injective,
- $\mathcal{R}(A_T)$ is dense in $L^2(0,\pi)$,
- A_T is linear, bounded and **compact** $\Rightarrow$ no continuous inverse.

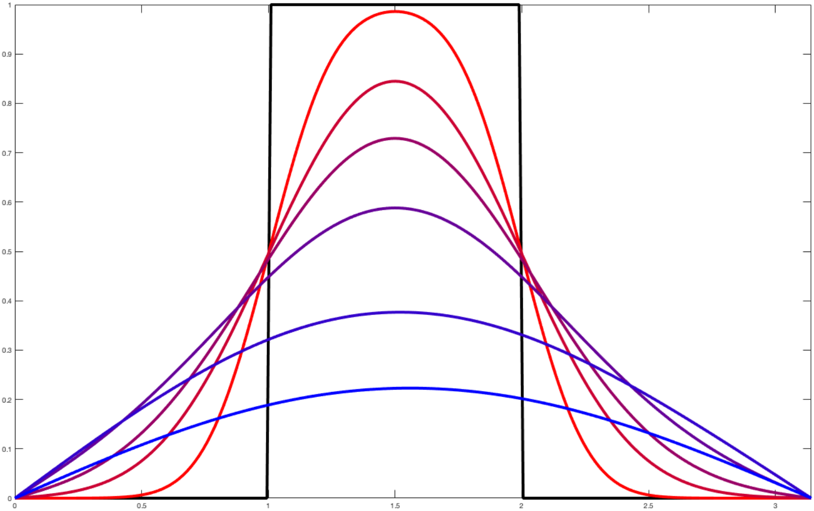
Inverse problem

We notice that, for every $s > 0$,

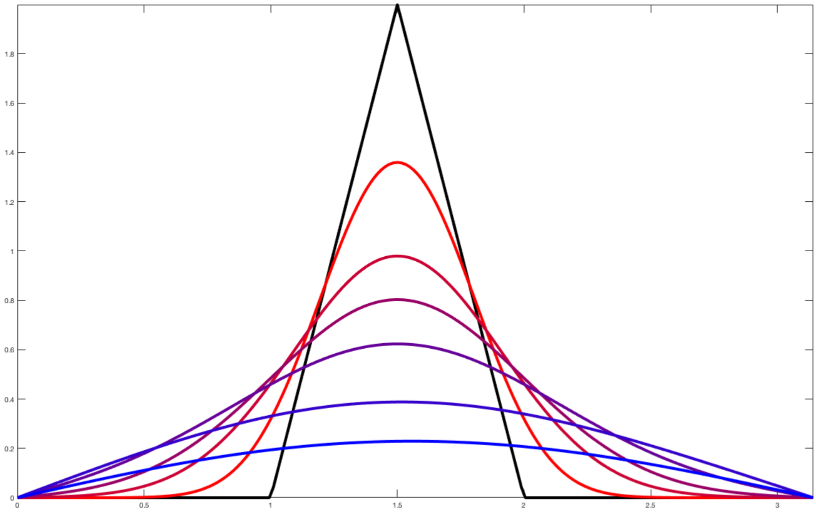
$$\begin{aligned}\|v\|_{H^s}^2 &= \sum_{j=1}^{\infty} j^{2s} e^{-2j^2 T} |\widehat{u}_j|^2 \\ &= T^{-s} \sum_{j=1}^{\infty} (j^2 T)^s e^{-2j^2 T} |\widehat{u}_j|^2 \\ &\leq C T^{-s} \|u\|_{L^2}^2\end{aligned}$$

and hence $\mathcal{R}(A_T) \subset \cap_{s>0} H^s$. However, noise is not smooth and hence $f_n = v(\cdot, T) + n \notin \mathcal{R}(A_T)$.

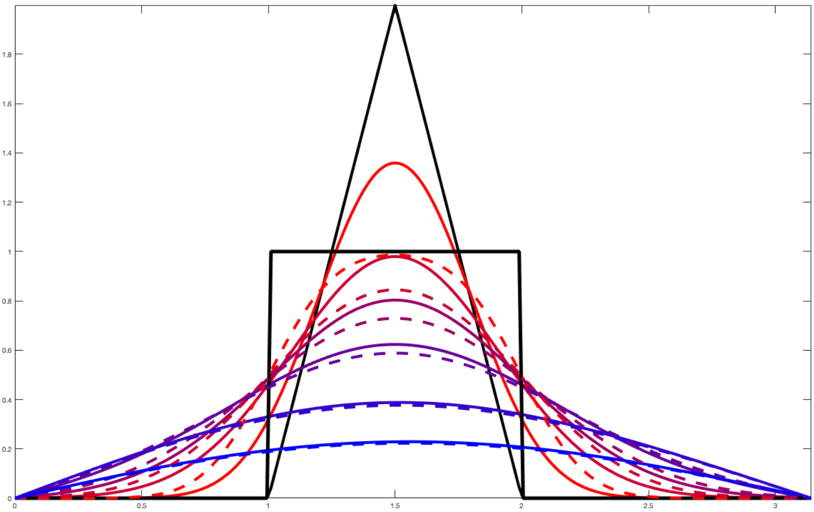
Heat distribution at $t = 0.02, 0.06, 0.1, 0.2, 0.5$ and $t = 1$



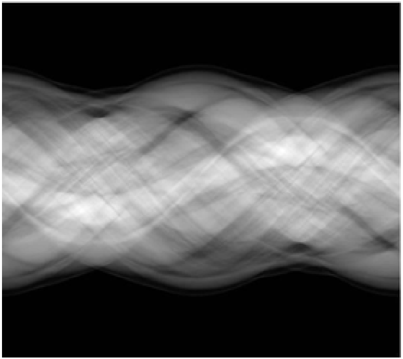
Another heat distribution at $t = 0.02, 0.06, 0.1, 0.2, 0.5$ and $t = 1$



Comparison of the two heat distributions



Examples: Computerised tomography (CT)



$$f(\theta, s) = (Au)(\theta, s) = \int_{x\cdot\theta=s} u(x)dx$$

The data are collected by rotating the X-ray source and detectors around the object

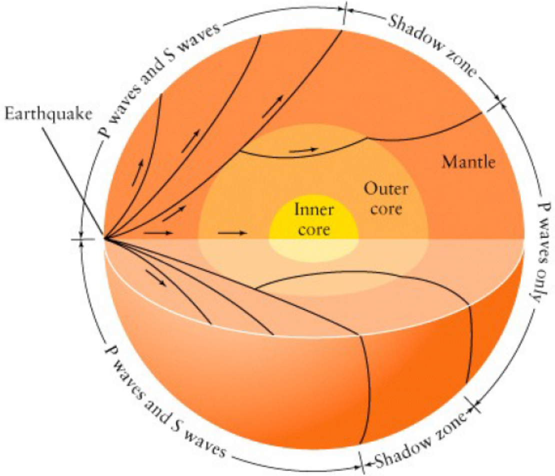
Video by Samuli Siltanen. For more videos see

`www.siltanen-research.net/IPexamples/xray_tomography`

Examples: Ozone layer tomography

Figure: Given star occultation measurements, what is the ozone profile?

Examples: Geodesic X-ray transform



$$f(\gamma) = (Au)(\gamma) = \int u(\gamma(t))dt$$

Examples: Photo-acoustic tomography

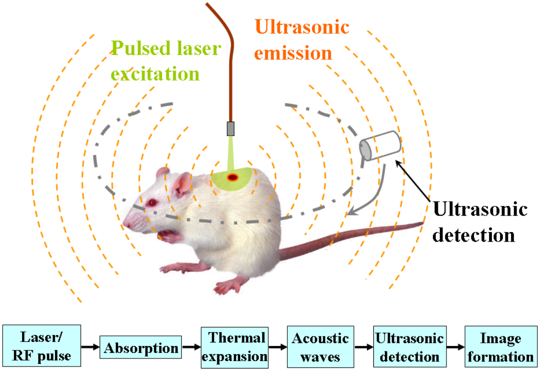


Figure: Photo-acoustic tomography is an example of a hybrid inverse problem that aims to combine high contrast and resolution of different imaging methods.

Examples: Photo-acoustic tomography

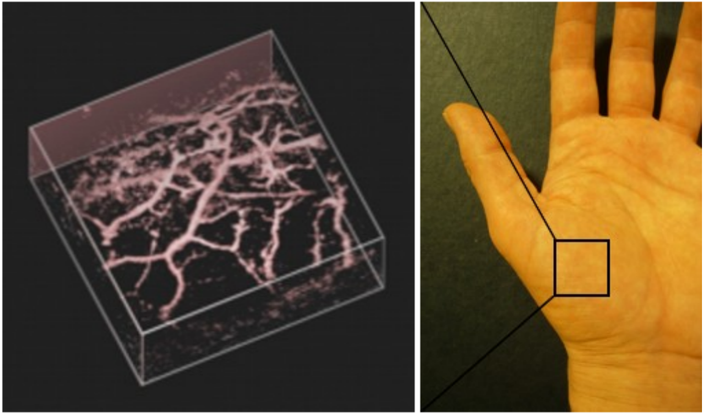


Figure: Images of superficial blood vessels. Image by UCL Photoacoustic Imaging Group

Examples: Hydraulic tomography

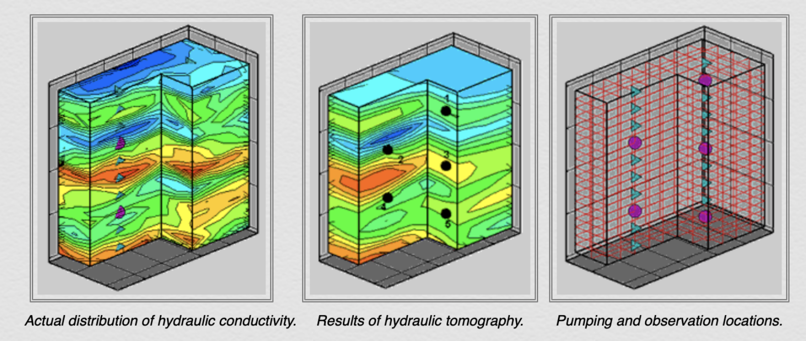


Figure: Hydraulic tomography: Measure water pressure in a groundwater reservoir to estimate the hydraulic conductivity (Image by Dr Jim Yeh; `http://tian.hwr.arizona.edu/research/HT/examples`)

$f_n = (p(x_i) : i = 1, ..., I)$, where

$$-\nabla \cdot \exp(u(x))\nabla p(x) = s(x) \qquad (x \in D) + \text{ b.c.}$$

NB. the map $u \mapsto f_n$ is non-linear

Outline of the course

1. Functional-analytic regularisation (ca. 12 lectures)
    - Regularisation theory in Hilbert spaces
        - singular value decomposition of compact operators, spectral filtering;
    - Regularisation theory in Banach spaces
        - variational methods, convex analysis, duality;
2. Bayesian inverse problems (ca. 12 lectures)
    - Uncertainty, statistics, and learning
        - Bayesian and conditional probability on separable Banach spaces, random fields
        - Bayesian inverse problems and the linear Gaussian case
    - Well-posedness of Bayesian inverse problems
        - existence, uniqueness, and stability of posterior measures
    - Algorithms for Bayesian inversion
        - Importance sampling, Markov chain Monte Carlo

Solving an inverse problem: Variational approach

We want to approximate u from a measurement

$$f_n = Au + n,$$

where $A : X \to Y$ is linear and bounded, X and Y are Banach spaces and $n \notin \mathcal{R}(A)$ is a perturbation such that $\|n\|_Y \leq \delta$.

One approach is to use the least squares method

$$\widetilde{u} = \arg\min_{u \in X} \left\{ \|Au - f_n\|_Y^2 \right\}.$$

Problem: Multiple minima (if A is not injective) and no stability with respect to the data f_n.

Solving an inverse problem: Variational approach

We want to approximate u from a measurement

$$f_n = Au + n,$$

where $A : X \to Y$ is linear and bounded, X and Y are Banach spaces and $n \notin \mathcal{R}(A)$ is a perturbation such that $\|n\|_Y \leq \delta$.

To regularise the problem we add a regularisation term and define

$$\widetilde{u} = \arg\min_{u \in X} \left\{ \|Au - f_n\|_Y^2 + \alpha \mathcal{R}(u) \right\}$$

Regularisation gives a stable approximate solution for the inverse problem. Parameter α depends on the magnitude δ of the perturbation n.

Solving an inverse problem: Variational approach

We want to approximate u from a measurement

$$f_n = Au + n,$$

where $A : X \to Y$ is linear and bounded, X and Y are Banach spaces and $n \notin \mathcal{R}(A)$ is a perturbation such that $\|n\|_Y \leq \delta$.

To regularise the problem we add a regularisation term and define

$$\widetilde{u} = \arg\min_{u \in X} \left\{ \|Au - f_n\|_Y^2 + \alpha \mathcal{R}(u) \right\}$$

Regularisation gives a stable approximate solution for the inverse problem. Parameter α depends on the magnitude δ of the perturbation n.

Problem: perturbation n does not have a statistical interpretation, the constraint $\|n\|_Y \leq \delta$ will not be satisfied for a random n.

Solving an inverse problem: Stochastic approach

Problem: perturbation n does not have a statistical interpretation, the constraint $\|n\|_Y \leq \delta$ will not be satisfied for a random n.
Problem: variational approach gives an estimate for the solution of

$$f_n = Au + n$$

Questions:

- how good is this estimate?
- statistical properties?
- uncertainties?

Solving an inverse problem: Stochastic approach

We consider a measurement model

$$f_n = \mathcal{A}(u) + n$$

u and n are uncertain $\overset{\text{[Cox, 1946]}}{\Longrightarrow}$ represent through random variables

Solving an inverse problem: Stochastic approach

We consider a measurement model

$$f_n = \mathcal{A}(u) + n$$

u and n are uncertain $\overset{\text{[Cox, 1946]}}{\Longrightarrow}$ represent through random variables

- U describing our a priori knowledge concerning the parameter

$$\mathbb{P}(U \in \cdot) \qquad \text{is called prior distribution}$$

- N describing our knowledge concerning the noise

Solving an inverse problem: Stochastic approach

We consider a measurement model

$$f_n = \mathcal{A}(u) + n$$

u and n are uncertain $\overset{\text{[Cox, 1946]}}{\Longrightarrow}$ represent through random variables

- U describing our a priori knowledge concerning the parameter

$$\mathbb{P}(U \in \cdot) \qquad \text{is called prior distribution}$$

- N describing our knowledge concerning the noise

obtaining the measurement is an event:

$$\{f_n = \mathcal{A}(U) + N\}$$

The solution is a probability distribution

"Learn" data by conditioning the parameter U on $\{f_n = \mathcal{A}(U) + N\}$, i.e., determine the posterior distribution:

$$\mathbb{P}(U \in \cdot | f_n = \mathcal{A}(U) + N)$$

The solution is a probability distribution

"Learn" data by conditioning the parameter U on $\{f_n = \mathcal{A}(U) + N\}$, i.e., determine the posterior distribution:

$$\mathbb{P}(U \in \cdot | f_n = \mathcal{A}(U) + N)$$

- posterior distribution describes our knowledge a posteriori to seeing the data
- posterior distribution is obtained through Bayes' formula

$\Rightarrow$ Bayesian inverse problem

The solution is a probability distribution

"Learn" data by conditioning the parameter U on $\{f_n = \mathcal{A}(U) + N\}$, i.e., determine the posterior distribution:

$$\mathbb{P}(U \in \cdot | f_n = \mathcal{A}(U) + N)$$

- posterior distribution describes our knowledge a posteriori to seeing the data
- posterior distribution is obtained through Bayes' formula

$\Rightarrow$ Bayesian inverse problem

Bayes' theorem (informal)

$$\pi_{\text{posterior}}(u|f_n) \propto L(f_n|u)\pi_{\text{prior}}(u),$$

$\pi_{\text{prior}}, \pi_{\text{posterior}}$ are prior/posterior densities, and L is the likelihood

The solution is a probability distribution

Finite dimensional Gaussian example

- $X = \mathbb{R}^d$ and $Y = \mathbb{R}^k$, $A : X \to Y$ linear,
- $N \sim \mathrm{N}(0, \mathrm{Id}_k)$ is white Gaussian noise
- We choose Gaussian prior $\mathbb{P}(U \in \cdot) = \mathrm{N}(0, \Sigma)$

Posterior has density

$$\pi_{\text{posterior}}(u|f_n) \propto \exp\left(-\frac{1}{2}\|f_n - Au\|^2_{\mathbb{R}^k} - \frac{1}{2}\|u\|^2_{\Sigma} \right)$$

The solution is a probability distribution

Finite dimensional Gaussian example

- $X = \mathbb{R}^d$ and $Y = \mathbb{R}^k$, $A : X \to Y$ linear,
- $N \sim \mathrm{N}(0, \mathrm{Id}_k)$ is white Gaussian noise
- We choose Gaussian prior $\mathbb{P}(U \in \cdot) = \mathrm{N}(0, \Sigma)$

Posterior has density

$$\pi_{\text{posterior}}(u|f_n) \propto \exp\left(-\frac{1}{2}\|f_n - Au\|^2_{\mathbb{R}^k} - \frac{1}{2}\|u\|^2_{\Sigma}\right)$$

- here: posterior has a closed form solution
- Optimising $\pi_{\text{posterior}}(u|f_n) \Rightarrow$ regularisation (with $\mathcal{R}(u) = \|u\|^2_{\Sigma}$)

Questions? Comments?