# Bayesian inverse problems

**Lent term 2019**

**Hanne Kekkonen**

January 17, 2019

UNIVERSITY OF CAMBRIDGE

# Practicalities

- Lectures take place on Tuesdays and Thursdays, 9-10am, MR14.
- Course materials (lecture notes, example sheets, etc.) will be provided at `http://www.damtp.cam.ac.uk/research/cia/teaching/bayesinvprob19.html`
- Three example sheets and example classes (details to follow).
- Revision class will be held in May.
- Written exam will be held on 10 June 2019, 9-11am.
- For further questions email hnk22@cam.ac.uk

# Example classes

Example classes will take place on the following dates:

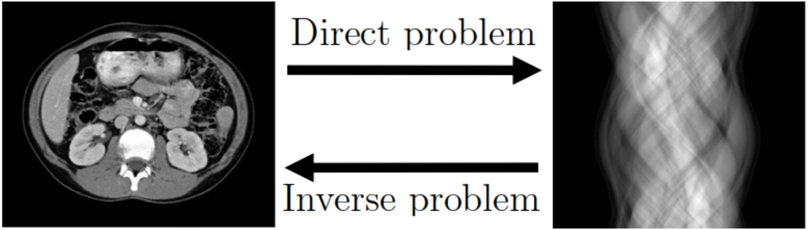
- Monday 11 February 2019, 2-3pm, MR4
- Monday 25 February 2019, 2-3pm, MR4
- Monday 11 March 2019, 2-3pm, MR4

Return solutions to two questions (specified in advance)

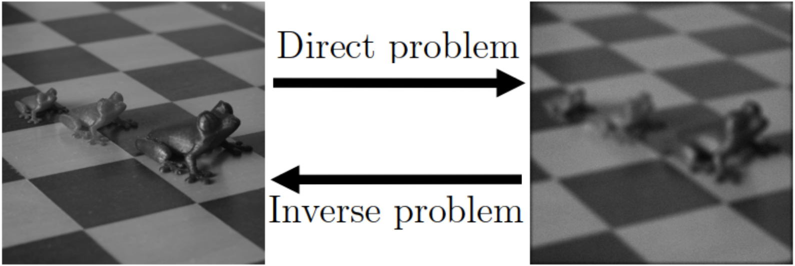
- Example sheets will be made available one week before the deadline.
- Hand in your answers by 4pm on the previous Friday.

# 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-posed problems were introduced 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

# We assume linear measurement model

We consider the linear measurement model

$$m_0 = Au$$

- The physical phenomena that relates the unknown and measurement is modelled by a linear operator $A : X \to Y$.
- The ideal measurement is given by $m_0 = Au$.

# We assume linear measurement model

We consider the linear measurement model

$$m_0 = Au$$

- The physical phenomena that relates the unknown and measurement is modelled by a linear operator $A : X \to Y$.
- The ideal measurement is given by $m_0 = Au$.
- We can only observe noisy measurements $m = Au + n$

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

# We assume linear measurement model

We consider the linear measurement model

$$m_0 = Au$$

- The physical phenomena that relates the unknown and measurement is modelled by a linear operator $A : X \to Y$.
- The ideal measurement is given by $m_0 = Au$.
- We can only observe noisy measurements $m = Au + n$

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

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

# Examples of problems caused by ill-posedness

I) $A : \mathbb{R}^d \to \mathcal{R}(A) \subsetneq \mathbb{R}^d$. Assume there is unique $A^{-1} : \mathcal{R}(A) \to \mathbb{R}^d$. Because of the noise $m \notin \mathcal{R}(A)$
$\Rightarrow$ We can't just invert $A$.

# Examples of problems caused by ill-posedness

I) $A : \mathbb{R}^d \to \mathcal{R}(A) \subsetneq \mathbb{R}^d$. Assume there is unique $A^{-1} : \mathcal{R}(A) \to \mathbb{R}^d$. Because of the noise $m \notin \mathcal{R}(A)$
$\Rightarrow$ We can't just invert $A$.

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

# Examples of problems caused by ill-posedness

I) $A: \mathbb{R}^d \to \mathcal{R}(A) \subsetneq \mathbb{R}^d$. Assume there is unique $A^{-1}: \mathcal{R}(A) \to \mathbb{R}^d$. Because of the noise $m \notin \mathcal{R}(A)$
$\Rightarrow$ We can't just invert $A$.

II) $A: \mathbb{R}^d \to \mathbb{R}^k$, $d > k$, 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 $\widetilde{u} = A^{-1}m = u + A^{-1}n$ is dominated by the noise.

# Examples: Deblurring (deconvolution)

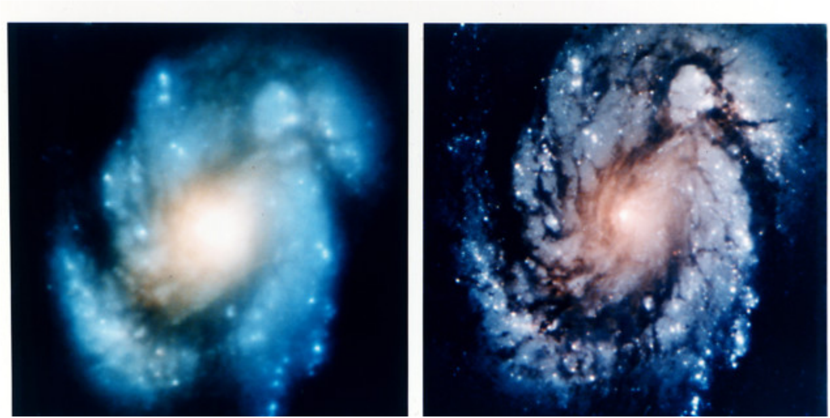


Figure: The Hubble space telescope had a flaw in its mirror which resulted in the images being blurred.

$$m_0(x) = (Au)(x) = \int_{\mathbb{R}^2} a(x-y)u(u)dy$$

# Signal deblurring for noiseless data

The noiseless data $m_0(t) = \int_{-\infty}^{\infty} a(t-s)u(s)ds$ has Fourier transform

$$\hat{m}_0(\xi) = \int_{-\infty}^{\infty} e^{-i\xi t} m_0(t) dt.$$

# Signal deblurring for noiseless data

The noiseless data $m_0(t) = \int_{-\infty}^{\infty} a(t-s)u(s)ds$ has Fourier transform

$$\hat{m}_0(\xi) = \int_{-\infty}^{\infty} e^{-i\xi t} m_0(t)dt.$$

The Convolution Theorem implies

$$\hat{m}_0(\xi) = \hat{a}(\xi)\hat{u}(\xi),$$

# Signal deblurring for noiseless data

The noiseless data $m_0(t) = \int_{-\infty}^{\infty} a(t-s)u(s)ds$ has Fourier transform

$$\hat{m}_0(\xi) = \int_{-\infty}^{\infty} e^{-i\xi t} m_0(t) dt.$$

The Convolution Theorem implies

$$\hat{m}_0(\xi) = \hat{a}(\xi)\hat{u}(\xi),$$

and hence by inverse Fourier transform

$$u(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{it\xi} \frac{\hat{m}_0(\xi)}{\hat{a}(\xi)} d\xi.$$

# Signal deblurring for noisy data

We can only observe noisy data and get

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

# Signal deblurring for noisy data

We can only observe noisy data and get

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

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

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

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.

# Examples: Heat distribution in an insulated rod

Consider the problem

$$\begin{aligned} v_t - v_{xx} &= 0 \quad \text{in } (0,\pi) \times \mathbb{R}_+ \\ v(0,\cdot) = v(\pi,\cdot) &= 0 \quad \text{on } \mathbb{R}_+ \\ v(\cdot,T) &= m_0 \quad \text{in } (0,\pi) \\ v(\cdot,0) &= u \quad \text{in } (0,\pi) \end{aligned}$$

# Examples: Heat distribution in an insulated rod

Consider the problem

$$\begin{aligned} v_t - v_{xx} &= 0 \quad \text{in } (0,\pi) \times \mathbb{R}_+ \\ v(0,\cdot) = v(\pi,\cdot) &= 0 \quad \text{on } \mathbb{R}_+ \\ v(\cdot,T) &= m_0 \quad \text{in } (0,\pi) \\ v(\cdot,0) &= u \quad \text{in } (0,\pi) \end{aligned}$$

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

$$
\begin{aligned}
v_t - v_{xx} &= 0 \quad \text{in } (0,\pi)\times\mathbb{R}_+ \\
v(0,\cdot) = v(\pi,\cdot) &= 0 \quad \text{on } \mathbb{R}_+ \\
v(\cdot,T) &= m_0 \quad \text{in } (0,\pi) \\
v(\cdot,0) &= u \quad \text{in } (0,\pi)
\end{aligned}
$$

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

$$A_T : u \mapsto v(\cdot, T) = m_0, \quad L^2(0,\pi) \to L^2(0,\pi)$$

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.

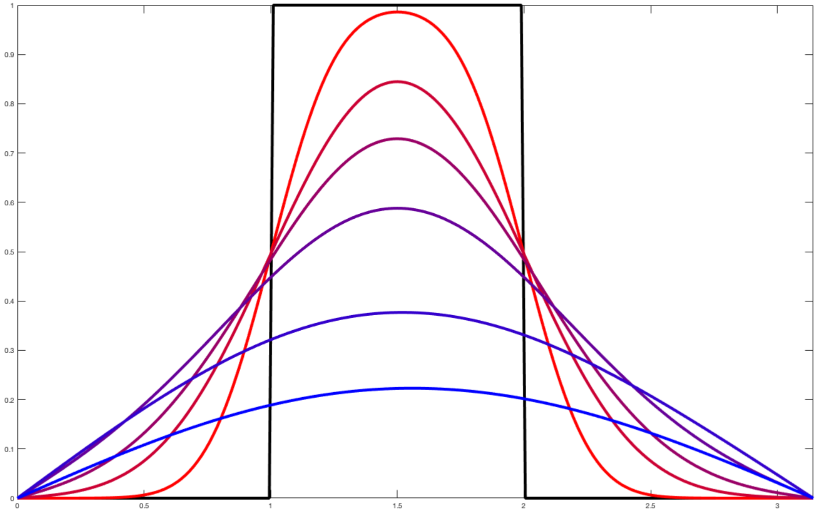
# Inverse problem

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

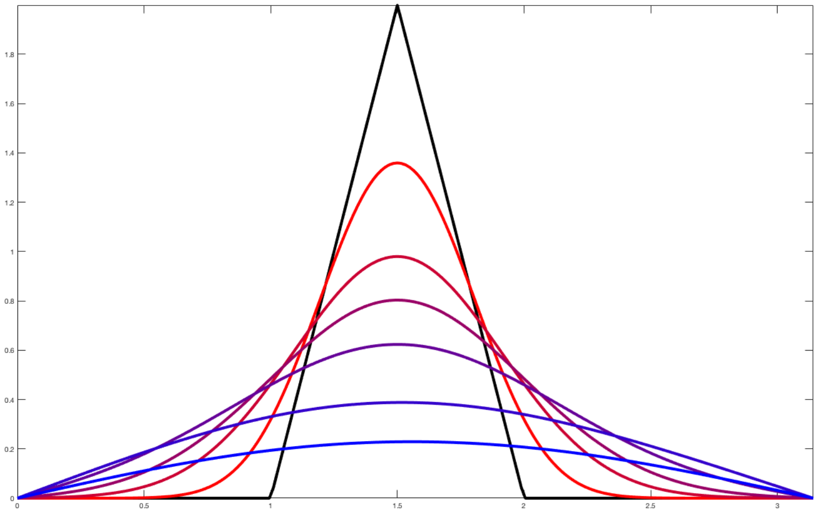
$$
\begin{aligned}
\|v\|_{H^s}^2 &= \sum_{j=1}^{\infty} j^{2s} e^{-2j^2 T} |\widehat{u}_j|^2 \\
&= T^{-s} \sum_{j=1}^{\infty} (j^2 T)^s e^{-2j^2 T} |\widehat{u}_j|^2 \\
&= C T^{-s} \|u\|_{L^2}^2
\end{aligned}
$$

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

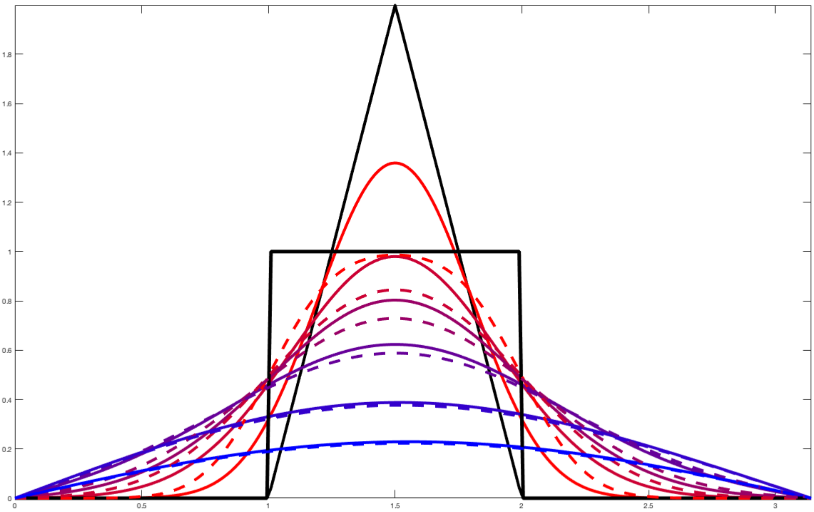
## Heat distribution at $t = 0.02, 0.06, 0.1, 0.2, 0.5$ and $t = 1$



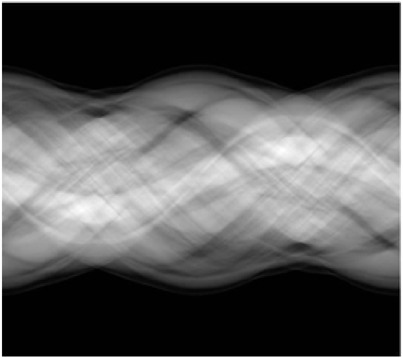
# Another heat distribution at $t = 0.02, 0.06, 0.1, 0.2, 0.5$ and $t = 1$



# Comparison of the two heat distributions



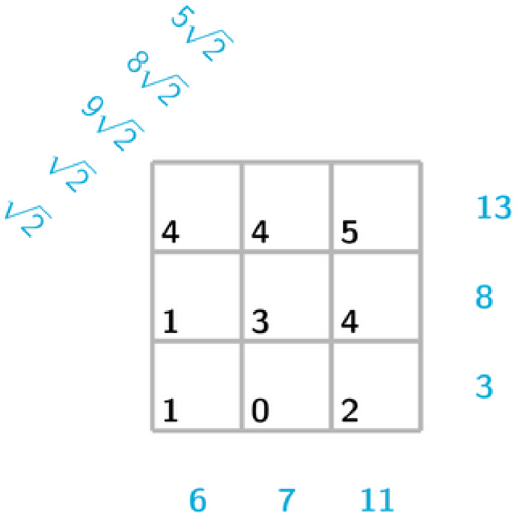
# Examples: Computerised tomography (CT)



$$m_0(\theta, s) = (Au)(\theta, s) = \int_{x \cdot \theta = s} u(x) dx$$

See videos at `www.siltanen-research.net/IPexamples/xray_tomography`

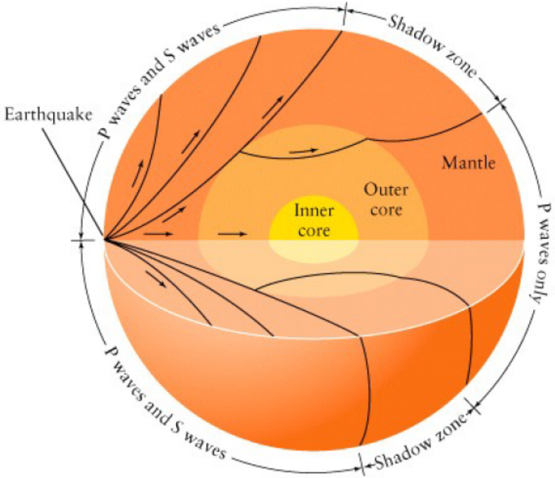
# Examples: Computerised tomography (CT)



# Examples: Ozone layer tomography

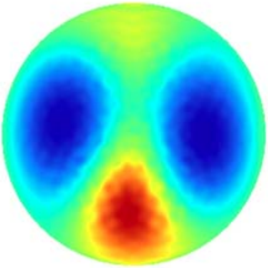
Figure: Given star occultation measurements, what is the ozone profile?

# Examples: Geodesic X-ray transform




$$m_0(\gamma) = (Au)(\gamma) = \int u(\gamma(t))dt$$

# Non-linear example: Electrical Impedance Tomography (EIT)



Applying electric voltages $f$ at the boundary leads to PDE

$$\nabla \cdot (\sigma \nabla v) = 0 \quad \text{in } \Omega \in \mathbb{R}^2$$
$$v|_{\partial\Omega} = f$$

Non-linear inverse problem: Recover conductivity $\sigma$ from boundary measurements $\Lambda_\sigma(f) = \sigma \frac{\partial v}{\partial \vec{n}}|_{\partial\Omega}$

# EIT in industry

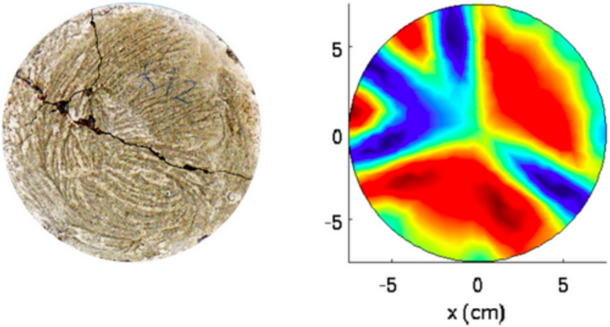



Figure: EIT and similar methods can be used for example to detect cracks in concrete and in industrial process monitoring.

# Examples: Asteroid lightcurve inversion

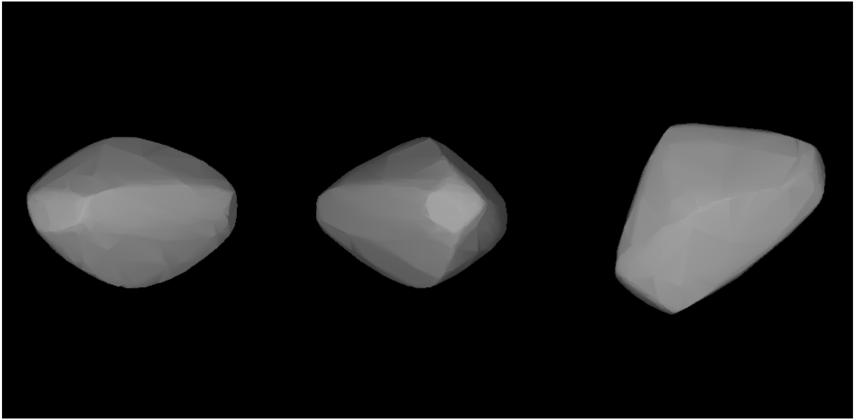


Figure: The lightcurve inversion technique is used to find an object's rotation period, its shape and spin-axis orientation

# Examples: Photo-acoustic tomography

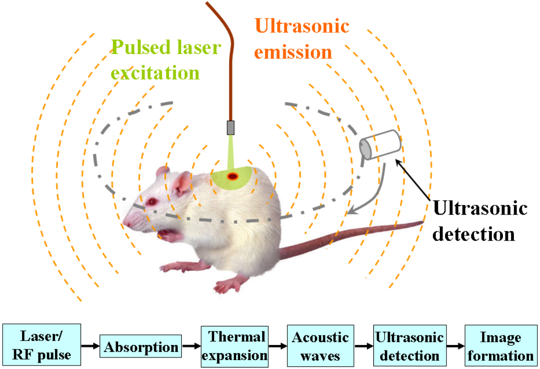



Figure: Photo-acoustic tomography is an example of a hybrid inverse problem that aims to combine high contrast and resolution of different imaging methods.

# Examples: Photo-acoustic tomography

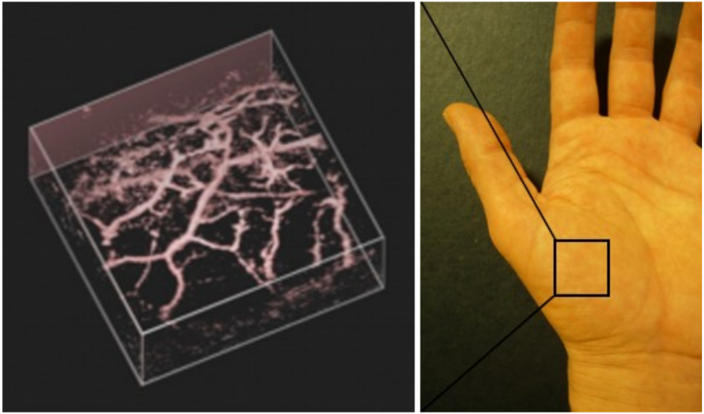


Figure: Images of superficial blood vessels. Image by UCL Photoacoustic Imaging Group

# Solving an inverse problem: Deterministic approach

We want to approximate $u$ from a measurement

$$m = Au + n,$$

where $A : X \to Y$ is linear and $n$ is noise.

One approach is to use the least squares method

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

Problem: Multiple minima and sensitive dependence on the data $m$.

# Solving an inverse problem: Deterministic approach

We want to approximate $u$ from a measurement

$$m = Au + n,$$

where $A : X \to Y$ is linear and $n$ is noise.

The problem is ill-posed so we add a regularising term and get

$$\widetilde{u} = \arg \min_{u \in E \subset X} \left\{ \|Au - m\|_Y^2 + \alpha \|u\|_E^2 \right\}$$

Regularisation gives a stable approximate solution for the inverse problem.

# Solving an inverse problem: Stochastic approach

We consider linear measurement model

$$m = Au + \eta,$$

where $m, u, \eta$ are now treated as random variables.

## Bayes' formula

$$\mathbb{P}(u \in A \,|\, m \in B) = \frac{\mathbb{P}(m \in B \,|\, u \in A)\mathbb{P}(u \in A)}{\mathbb{P}(m \in B)}$$

# The solution is a probability distribution

## Finite dimensional Gaussian example

- $X = \mathbb{R}^d$ and $Y = \mathbb{R}^k$
- $\eta$ is white Gaussian noise
- We choose Gaussian prior

Posterior has density

$$\pi^m(u) = \pi(u \mid m) \propto \exp\left( -\frac{1}{2}\|m - Au\|^2_{\mathbb{R}^k} - \frac{1}{2}\|u\|^2_{\Sigma} \right)$$

We can use the mean of the posterior as a point estimator but having the whole posterior allows uncertainty quantification.

# Bayesian analysis has many applications

Studying the whole posterior distribution instead of just a point estimate offers us more information.

## Uncertainty quantification

- Confidence and credible sets
- E.g. Weather and climate predictions

## Using the whole posterior

- Geological sensing
- Bayesian search theory

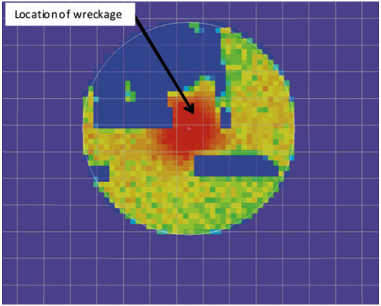



Figure: *Search for the wreckage of Air France flight AF 447, Stone et al.*