MEDICAL IMAGING

Lecture Skriptum

Benjamin Bergmann

2026

Contents

Inverse Problems

What is an Inverse Problem?

There exit a “Forward Problem” which estimate the effect from the cause and then there is inverse Problem which estimates the cause from the effect. In the medical context that would be finding the cause illness given from a certain syntom/effect. Typically, the forward problem is ”easy” and well described. The challenge here is: We need to solve the inverse problem given only the observed effect of the forward problem.

As an Example from the real world: forward problem: The street becomes wet when it rains. backward problem would be: We observe that the street is wet. Why?

There are multiple different causes: • Rain • Fog • Cleaning

And this can be already problematic as we have multiple different options for what the cause could be.

Example: Computed Tomography

Forward Problem X-ray emitter and detector rotating around the body. Detectors measure the number of photons passing through the body and hitting the detector

Inverse Problem Reconstruct the interior of the body from the measured detector signals.

Note that a CT Scan can be very large in file size. A scan from shoulder to belt line is already 18GB of data for just a single scan. So we bascially have y and we want to get to x

Example: Deconvolution

Forward Problem Observe a blurred image

$$f = k * u$$

on a domain

$$\Omega \subset \mathbb{R}^2$$

.

Inverse Problem Estimate the sharp image

$$u : \Omega \to \mathbb{R}$$

given the blur kernel

$$k : \Omega \times \Omega \to \mathbb{R}_+$$

On of the oldest clasical methods to do that is the Wiener Filter. Deconvolution is linked to Fourier F:

$$f = k * u$$

$$F(f) = F(k) \odot F(u)$$

If we wanna do the inverse:

$$F^{-1}\{F(f)\} = F^{-1}\{F(k) \odot F(u)\} = f$$

where $\odot$ is a pointwise multiplication

So a estimate $\hat{u}$ would be

$$\hat{u} = F^{-1}\left(\frac{F(f)}{F(k)}\right)$$

The only problem here is when we have 0 frequencies in the kernel. The Wiener Filtering introduces

$$\hat{u} = F^{-1}\left(\frac{F(f)}{I\sigma^2 F(k)}\right)$$

What is an Inverse Problem? (formal)

> **Definition 1** (Inverse Problem) . Given a matrix $A \in \mathbb{R}^{m\times n}$ and a vector $x \in \mathbb{R}^n$ the forward problem is $y = Ax \in \mathbb{R}^m$
>
> The inverse problem is: Given A and y, estimate x.

Vector Space

> **Definition 2** (Vector Space) . A non-empty set V is a vector space over a field $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ if there are operations of vector addition: $+ : V \times V \to V$ and scalar multiplication: $\cdot : \mathbb{F} \times V \to V$ satisfying the following axioms:
>
> **Vector addition**
> 1. $u + v \in V \quad \forall u, v \in V$
> 2. $u + v = v + u$
> 3. $(u + v) + w = u + (v + w) \quad \forall u, v, w \in V$
> 4. $\exists 0 \in V : u + 0 = u \quad \forall u \in V$
> 5. $\forall u \in V : \exists {-w} : u + (-w) = 0$
>
> **Scalar multiplication**
> 1. $av \in V \quad \forall a \in \mathbb{F}, \forall v \in V$
> 2. $(ab)v = a(bv) \quad \forall a, b \in \mathbb{F}, v \in V$
> 3. $a(u + v) = au + av \quad \forall a \in \mathbb{F}, \forall u, v \in V$
> 4. $(a + b)v = av + bv \quad \forall a, b \in \mathbb{F}, \forall v \in V$
> 5. $\exists 1 \in \mathbb{F} : 1 * u = u \quad \forall u \in V$

Vector Space Examples

- $\mathbb{R}^n = \left\{(x_1, ..., x_n)^T : x_1, ..., x_n \in \mathbb{R}\right\}$
- $\mathcal{C}(\mathbb{R}^n, \mathbb{R})$ set of function $f : \mathbb{R}^n \to \mathbb{R}$ that are continuous
- $\mathcal{C}^1(\mathbb{R}^n, \mathbb{R})$ set of function $f : \mathbb{R}^n \to \mathbb{R}$ that are continuous and once continuously differentiable
- $L^2(\mathbb{R}^n, \mathbb{R}) = \left\{f : \mathbb{R}^n \to \mathbb{R} : \int_{\mathbb{R}^n} |f(x)|^2 \, dx < \infty\right\}$ Lebesgue space
- $H^1(\mathbb{R}^n, \mathbb{R}) = \left\{f \in L^2(\mathbb{R}^n, \mathbb{R}) : \int_{\mathbb{R}^n} |f'(x)|^2 \, dx < \infty\right\}$ Sobolev space ($p = 2$), Hilbert space

Inverse Problem

Definition 3 (Inverse Problem) . Let X, Y be vector spaces and $A : X \to Y$. The forward problem is defined as $y = Ax$ for any $x \in X$. The inverse problem is to find $x \in X$ such that $Ax = y$ for any $y \in Y$.

So we want to get $A^{-1}(y) = \hat{x}$

Well-Posedness (Hadamard)

We can now start to categorize inverse problems:

Definition 4 (Well-Posedness) . The inverse problem $Ax = y$ is well-posed if:

1. **Existence:** a solution exists (EXISTENCE)
2. **Uniqueness:** the solution is unique (UNIQUENESS)
3. **Stability:** the solution depends continuously on the data (STABILITY)

If one condition fails, the problem is ill-posed.

Well-Posedness Example

Example 1)

Is this example well posed? Let $X, Y = \mathbb{R}$ and $A : \mathbb{R} \to \mathbb{R}, x \to x^2$

Answer:

- Existence: for $y = -1$ no solution exists (if it would be $\mathbb{R}^+$ it would be okay)
- Uniqueness: for $y = 1$, $x = \pm 1$ which is not unique
- Stability: yes, since A is continuous

Example 2)

Let $X, Y = \mathbb{R}^2$ and $A = \begin{pmatrix} 2 & 3 \\ 1 & 2 \end{pmatrix} \in \mathbb{R}^{2 \times 2}$. Is the inverse problem $Ax = y$ for $y \in Y$ well-posed?

1. **EXISTENCE:** $\exists A^{-1}$? Since $\det(A) = 4 - 3 = 1 \neq 0$, the matrix is invertible.
2. **UNIQUENESS:** Yes, because $\det(A) \neq 0$.
3. **STABILITY:** Yes, as A^{-1} is continuous.

Inner Product

Definition 5 (Inner Product) . An inner product on a vector space Y over a $\mathbb{F}$ is a map

$$\langle ., . \rangle : Y \times Y \to \mathbb{F}$$

with the following properties:

1. Symmetry: $\langle x, y \rangle = \overline{\langle y, x \rangle} \quad x, y \in Y$
2. Additivity: $\langle x, y + z \rangle = \langle x, y \rangle + \langle x, z \rangle \quad x, y, z \in Y$
3. Homogeneity: $\langle \lambda x, y \rangle = \lambda \langle x, y \rangle \quad x, y \in Y \quad \lambda \in \mathbb{R}$
4. Positivity: $\langle x, x \rangle \geq 0$ and $\langle x, x \rangle = 0 \iff x = 0$

Vector Norm

Definition 6 (Inner Product) . A vector norm is a vector space Y over a field F is a map $\|.\| : Y \to \mathbb{R}$ with:

1. **NON-NEGATIVITY** $\|x\| \geq 0 \quad \forall x \in V, \|x\| = 0 \Leftrightarrow x = 0$
2. **POSITIVE HOMOGENEITY** $\|\lambda x\| = |\lambda| \; \|x\| \; \forall x \in Y, \lambda \in \mathbb{F}$
3. **TRIANGLE INEQUALITY** $\|x + y\| \leq \|x\| + \|y\| \; x, y \in V$

Example: l_p-norm

$$\|x\|_p = \sqrt[p]{\sum_{i=1}^{n} |x_i|^p} \quad x \in X \subset \mathbb{R}^n$$

Definition: Matrix Norm

Definition 7 (Inner Product) . Let $\|\cdot\|_a$ and $\|\cdot\|_b$ be vector norms on $\mathbb{R}^n$ and $\mathbb{R}^m$, respectively. Given a matrix $A \in \mathbb{R}^{m \times n}$, the **induced matrix norm** $\|A\|_{a,b}$ is defined as:

$$\|A\|_{a,b} = \max_{x \in \mathbb{R}^n : \|x\|_a \leq 1} \|Ax\|_b = \sup_{\{x \in \mathbb{R}^n \setminus \{0\}\}} \frac{\|Ax\|_b}{\|x\|_a}$$

$$\|Ax\|_b \leq \|A\|_{ab} \|x\|_a$$

Matrix Norm Examples
- If $a, b = 2$: $\|A\|_{2,2} = \|A\|_2 = \sigma_{\max}(A) = \sqrt{\lambda_{\max}(A^T A)}$
- If $a, b = 1$: $\|A\|_{1,1} = \|A\|_1 = \max_j \sum_i |A_{ij}|$
- If $a, b = \infty$: $\|A\|_{\infty} = \max_i \sum_j |A_{ij}|$

Injection, Surjection, Bijection

These properties of mappings $A : X \to Y$ are defined as

- **Injection:** $A : X \to Y$ is injective if $Ax_1 = Ax_2 \Rightarrow x_1 = x_2$.
- **Surjection:** $A : X \to Y$ is surjective if $\forall y \in Y, \exists x \in X : Ax = y$.
- **Bijection:** $A : X \to Y$ is bijective if it is both injective and surjective. $\forall y \in Y, \exists! x \in X : Ax = y \Leftrightarrow \exists A^{-1} : x = A^{-1} y$.

Null Space and Range

Let $A : X \to Y$ where X, Y are vector spaces.
- **Nullspace of A:** $N(A) = \{x \in X : Ax = 0\}$
- **Range space of A:** $R(A) = \{Ax \in Y : x \in X\}$

Connection to Hadamard's Definition

- **Existence** $\Leftrightarrow$ Surjection $\Leftrightarrow$ $R(A) = Y$
- **Uniqueness** $\Leftrightarrow$ Injection $\Leftrightarrow$ $N(A) = \{0\}$
- **Existence & Uniqueness** $\Leftrightarrow$ Bijection

Definition of the Linear Inverse Problem

Given $A : X \to Y$ and observation $y \in Y$ the inverse problem is called linear if A is linear which means that $A(\alpha x_1 + \beta x_2) = \alpha A(x_1) + \beta A(x_2)$

Example:

A… is the Radon transform

$$(Ax)_i = y_i = \int_{\Gamma_i} x(s) ds$$

$$A(\hat{x}) = A(\lambda_1 \cdot x_1 + \lambda_2 x_2) = \hat{y}_i = \int_{\Gamma_i} \hat{x}(s) ds = \int_{\Gamma_i} \lambda_1 x_1(s) + \lambda_2 \cdot x_2(s) ds$$
$$= \lambda_1 \underbrace{\int_{\Gamma_i} x_1(s) ds}_{y_i^1} + \lambda_2 \underbrace{\int_{\Gamma_i} x_2(s) ds}_{y_i^2} = \lambda_1 y_i^1 + \lambda_2 y_i^2 = \lambda_1 A(x_1)_i + \lambda_2 A(x_2)_i$$

Nullspace of linear A $\Rightarrow \{0\} \in \mathcal{N}(A)$

Decomposition of Square Matrices

Let $A \in \mathbb{R}^{n \times n}$, recall Eigenvalues λ_i and Eigenvectors v_i:

$$Av_i = \lambda_i v_i \quad \text{for } i = 1, ..., n$$

$$\det(A - \lambda I) = 0$$

If v_i are linearly independent: $Av_i = \lambda_i v_i \Rightarrow AQ = Q\Lambda \Rightarrow A = Q\Lambda Q^{-1}$ Where $Q = (v_1, ..., v_n)$.

Remark: If A is hermitian $\Leftrightarrow A^* = A$, we have that all λ_i are real & v_i are orthonormal.

$$v_i^T v_j = 0 \quad \text{for } i \neq j$$

$$A = Q\Lambda Q^T$$

Singular Value Decomposition

Let $X = \mathbb{R}^n, Y = \mathbb{R}^m$ be an inverse problem $Ax = y$ with a $A \in \mathbb{R}^{m \times n}$. The Goal:

$$A = U\Lambda V^T$$

- $U \in \mathbb{R}^{m \times p}, \Lambda \in \mathbb{R}^{p \times p}, V \in \mathbb{R}^{p \times n}$
- p is the number of non-zero singular values $\sigma_1 \geq \sigma_2 \geq ... \geq \sigma_p > 0$.

Link between SVD and Eigendecompostion

$$A \in \mathbb{R}^{m \times n}$$

$$\begin{cases} (1) & Ax = y \\ (2) & A^T \hat{x} = \hat{y} \end{cases} \Leftrightarrow \underbrace{\begin{pmatrix} 0 & A \\ A^T & 0 \end{pmatrix}}_{B \in \mathbb{R}^{(m+n) \times (m+n)}} \cdot \begin{pmatrix} \hat{x} \\ x \end{pmatrix} = \begin{pmatrix} y \\ \hat{y} \end{pmatrix}$$

$B = B^T : \quad Bw_i = \lambda_i w_i$

$$\begin{pmatrix} 0 & A \\ A^T & 0 \end{pmatrix} \begin{pmatrix} u_i \\ v_i \end{pmatrix} = \lambda_i \begin{pmatrix} u_i \\ v_i \end{pmatrix} \Leftrightarrow \begin{cases} \text{1st:} & Av_i = \lambda_i u_i \\ \text{2nd:} & A^T u_i = \lambda_i v_i \end{cases}$$

1st:

$$\lambda_i A v_i = \lambda_i^2 u_i$$

$$A(\lambda_i v_i) = \lambda_i^2 u_i$$

$$AA^T u_i = \lambda_i^2 u_i$$

$$U = (u_1 \mid \dots \mid u_m)$$

2nd:

$$A^T(\lambda_i u_i) = \lambda_i^2 v_i$$

$$A^T A v_i = \lambda_i^2 v_i$$

$$V = (v_1 \mid \dots \mid v_n)$$

Least Squares (m > n)

$Ax = y \quad A \in \mathbb{R}^{m\times n} \quad m > n$ overdetermined system

$$e_i = a_i^T x - y_i$$

Idea: minimize the squared error

$$\hat{x} = \arg\min E(x) := \frac{1}{2}\sum_{i=1}^{m} \left(a_i^T x - y_i\right)^2 = \frac{1}{2}\, \|Ax - y\|_2^2 = \frac{1}{2}\, \|e\|_2^2$$

where $e = Ax - y$

How do we solve this optimization problem?

$$\nabla E(x) = 0 = \frac{\partial e}{\partial x}\frac{\partial E}{\partial e} = \frac{\partial e}{\partial x}\frac{1}{2}2e = A^T e = A^T(Ax - y) = 0$$

$$\nabla E(x) \in \mathbb{R}^n$$

Least squares solution

$$A \in \mathbb{R}^{m\times n}$$

$$\nabla E = A^T(Ax - y) = 0$$

$$(A^T A)x = A^T y$$

$$x = \left(A^T A\right)^{-1} A^T y$$

Example: 2×2 CT reconstruction

$$x \in \mathbb{R}^4 \quad y \in \mathbb{R}^5$$

$$Ax = y$$

$$\begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \end{pmatrix}$$

Solving Inverse Problems ($p = n > m$)

Let $Ax = y$ with $A \in \mathbb{R}^{m\times n}$

Remark: Since $n > m$, $\left(A^T A\right)^{-1}$ does not exist. This is an **underdetermined system**. Multiple solutions exactly solve $Ax = y$. We pick one using a priori knowledge:

$$\min_x \frac{1}{2} \|x\|_2^2 \quad \text{s.t.} \quad Ax = y$$

Recap: Lagrange Multipliers To solve $\min E(x)$ subject to $C(x) = 0$

Define Lagrangian: $\mathcal{L}(x, \tau) = E(x) + \langle C(x), \tau \rangle$

Find solution by $\nabla \mathcal{L}(x, \tau) = 0$

$$\begin{cases} \frac{\partial}{\partial x} \mathcal{L} = \frac{\partial E}{\partial x} + \frac{\partial C}{\partial x} \tau = 0 \\ \frac{\partial}{\partial \tau} \mathcal{L} = C(x) = 0 \end{cases}$$

Minimum Length Solution

Find x s.t. $Ax = y$ and $\|x\|_2^2 \to \min$

$$E(x) = \frac{1}{2} \|x\|_2^2 \quad C(x) = y - Ax = 0 \Leftrightarrow h(x, \tau) = \frac{1}{2} \|x\|_2^2 + \langle y - Ax, \tau \rangle$$

$$\frac{\partial}{\partial x} \mathcal{L} = x - A^T \tau = 0 \Leftrightarrow x = A^T \tau$$

$$\frac{\partial}{\partial \tau} \mathcal{L} = y - Ax = 0 \Leftrightarrow y = Ax = A\left(A^T \tau\right) = \left(AA^T\right) \tau \Leftrightarrow \tau = \left(AA^T\right)^{-1} y$$

$$x = A^T \left(AA^T\right)^{-1} y$$

Generalized Inverse

Let $X = \mathbb{R}^n$, $Y = \mathbb{R}^m$ and the inverse problem $Ax = y$ with $A \in \mathbb{R}^{m \times n}$.

Define the **generalized inverse** as:

$$A_g^{-1} = \left(U_p \Lambda_p V_p^T\right)^{-1} = \left(V_p^T\right)^{-1} \Lambda_p^{-1} U_p^{-1} = V_p \Lambda_p^{-1} U_p^T$$

Check if GI computes Exact, LS, ML:

I. $p = m = n$**:**

$$A_g^{-1} = V_p \Lambda_p^{-1} U_p^T \quad | \cdot A = U_p \Lambda_p V_p^T$$

$$A_g^{-1} A = V_p \Lambda_p^{-1} \underbrace{U_p^T U_p}_{I} \Lambda_p V_p^T = V_p \underbrace{\Lambda_p^{-1} \Lambda_p}_{I} V_p^T = I$$

II. $p = m > n$**:**

$$\begin{aligned}
x &= \left(A^T A\right)^{-1} A^T y \\
&= \left(\left(U_p \Lambda_p V_p^T\right)^T \left(U_p \Lambda_p V_p^T\right)\right)^{-1} \left(U_p \Lambda_p V_p^T\right)^T y \\
&= \left(V_p \Lambda_p \underbrace{U_p^T U_p}_{\text{Id}} \Lambda_p V_p^T\right)^{-1} \left(V_p \Lambda_p U_p^T\right) y \\
&= \left(V_p \Lambda_p^2 V_p^T\right)^{-1} V_p \Lambda_p U_p^T y \\
&= V_p \Lambda_p^{-2} \underbrace{V_p^T V_p}_{\text{Id}} \Lambda_p U_p^T y \\
&= V_p \Lambda_p^{-1} U_p^T y = A_g^{-1} y
\end{aligned}$$

III. $p = m < n$**:**

$$\begin{aligned}
x &= A^T \left(A A^T\right)^{-1} y \\
&= \left(V_p \Lambda_p U_p^T\right)\left(U_p \Lambda_p V_p^T V_p \Lambda_p U_p^T\right)^{-1} y \\
&= \left(V_p \Lambda_p U_p^T\right)\left(U_p \Lambda_p^2 U_p^T\right)^{-1} y \\
&= V_p \Lambda_p \underbrace{U_p^T U_p} \Lambda_p^{-2} U_p^T y \\
&= V_p \Lambda_p^{-1} U_p^T y = A_g^{-1} y
\end{aligned}$$

IV. $0 < p < \min(m, n)$**:** However, A_g^{-1} still exists. It computes a solution that **interpolates between LS & ML solutions**.

Regularization

Consider polynomial regression

$$p(a) = \sum_{i=1}^{n} x_i a^{i-1} = x_1 \cdot 1 + x_2 \cdot a + \ldots + x_n a^{n-1}$$

Where x represents the **coefficients of the polynomial**.

$$\Leftrightarrow Ax = \begin{pmatrix} 1 & a_1 & a_1^2 & \ldots & a_1^{n-1} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & a_m & a_m^2 & \ldots & a_m^{n-1} \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix}$$

The matrix A has dimensions $[m \times n]$.

How to choose n?
- Manually
- Very large + regularization

Incorporating Prior Knowledge

Least squares problem + regularization:

$$\hat{x} = \arg\min_x \frac{1}{2} \|Ax - y\|_2^2 + R(x)$$

Example: $R(x) = \frac{\lambda}{2} \|x\|_2^2$ (Tikhonov regularization, aka weight decay)

Derivation:

$$\frac{1}{2}\,\|Ax-y\|^2+\frac{\lambda}{2}\,\|x\|_2^2\to\min$$

$$\frac{1}{2}\cdot 2\cdot A^T(Ax-y)+\frac{\lambda}{2}\cdot 2x=0$$

$$A^TAx+\lambda x=A^Ty\Leftrightarrow x=\left(A^TA+\lambda Id\right)^{-1}A^Ty$$

Regularization Types and Intuition

Name	$R(x)$	Intuition
Tikhonov	$\lambda\,\|Gx\|_2^2$	Existence of Inverse
L^2	$\lambda\,\|x\|_2^2$ ($G = Id$)	Minimum length/norm
H^1	$\lambda\,\|\nabla x\|_2^2$ ($G = \nabla$)	Smooth gradients
L^1	$\lambda\,\|x\|_1$	Sparse solutions
Total variation (TV)	$\lambda\,\|\nabla x\|_1$	Sparse gradients (piece-wise constant solutions)

The Proximal Mapping

1. Projection onto a set S

$$\mathrm{proj}_S(x)=\arg\min_{y\in S}\frac{1}{2}\,\|x-y\|_2^2$$

2. Proximal mapping of a function $g(x)$

$$\mathrm{prox}_g(x)=\arg\min_{y}\frac{1}{2}\,\|x-y\|_2^2+g(y)$$

If we define $g(y)$ as the indicator function:

$$g(y)=\begin{cases}0 & \text{if } y\in S\\ \infty & \text{else}\end{cases}$$

Example: $g(x)=|x|$ To find the proximal mapping for the absolute value (L1 norm), we solve:

$$\mathrm{prox}_{|\cdot|}(x)=\arg\min_{y}\frac{1}{2}(x-y)^2+|y|$$

The subdifferential of $|x|$ is:

$$\frac{d}{dx}\,|x|=\begin{cases}1 & x>0\\ [-1,1] & x=0\\ -1 & x<0\end{cases}$$

To minimize, we set the subgradient to zero:

$$x - y + \partial g(y) \text{ contents } 0$$

- **Case $y > 0$:** $-(x - y) + 1 = 0 \Rightarrow y = x - 1 > 0 \Rightarrow x > 1$
- **Case $y < 0$:** $-x + y - 1 = 0 \Rightarrow y = x + 1 < 0 \Rightarrow x < -1$
- **Case $y = 0$:** $-x + 0 + [-1, 1]$ contents $0 \Rightarrow x \in [-1, 1]$

Thus, the Soft Thresholding operator is:

$$\operatorname{prox}_{|\cdot|}(x) = \begin{cases} x - 1 & \text{if } x > 1 \\ x + 1 & \text{if } x < -1 \\ 0 & \text{else} \end{cases}$$

Regularization IV: A Probabilistic Perspective

Assume observed measurements y follow a Gaussian distribution:

$$y \sim \mathcal{N}(Ax, \Sigma) \Longleftrightarrow p(y|x) = |2\pi\Sigma|^{-\frac{1}{2}} \exp\left(-\frac{1}{2} \|Ax - y\|^2_{\Sigma^{-1}}\right)$$

Moreover, we assume the solution (or its gradients) follows a Gaussian prior:

$$\nabla x \sim \mathcal{N}(0, \eta \text{ Id}) \Longleftrightarrow p(x) = |2\pi\eta I|^{-\frac{1}{2}} \exp\left(-\frac{1}{2}\eta \|x\|^2\right)$$

Using Bayes' Rule to find the posterior distribution:

$$p(x|y) = \frac{p(y|x) \cdot p(x)}{p(y)}$$

Taking the logarithm:

$$\log p(x|y) = \log p(y|x) + \log p(x) - \log p(y)$$

$$\log p(x|y) = -\frac{1}{2} \|Ax - y\|^2_{\Sigma^{-1}} - \log Z_1 - \frac{1}{2\eta} \|x\|^2 - \log Z_2 - \log p(y)$$

Since Z_1, Z_2, and $p(y)$ are constants that do not depend on x:

$$\max_x \log p(x|y) = \max_x -\frac{1}{2} \|Ax - y\|_{\Sigma}^{-} 1^2 - \frac{1}{2\eta} \|x\|^2$$

$$\min_x - \log p(x|y) = \min_x \underbrace{\frac{1}{2} \|Ax - y\|^2_{\Sigma^{-1}}}_{D(x,y) \;\; \text{(Data Fidelity)}} + \underbrace{\frac{1}{2\eta} \|x\|^2}_{R(x) \;\; \text{(Regularizer)}}$$

Conclusion: The variational formulation of inverse problems corresponds to the Maximum A Posteriori (MAP) estimation.

X-rays and CT

Discovery of X-rays

- In 1895 Wilhelm Röntgen discovered "rays of mysterious origin", later called X-rays.
- On 22.12.1895 the first radiograph of the hand of Röntgen's wife was produced.
- This immediate medical application marks the birth of medical imaging.

Nature and Properties of X-rays

- X-rays are electromagnetic waves.
- They are a form of ionizing radiation, i.e. radiation with enough energy to eject electrons from atoms.

Ionizing Radiation

Two main forms:

1. **Particulate radiation** Subatomic particles (electrons, protons, neutrons) with sufficient kinetic energy.
2. **Electromagnetic radiation** Acts as wave or particle (photon).

EM radiation is ionizing if photon energy exceeds the hydrogen binding energy: $E > 13.6eV$

Relations: $E = h\nu$ $\lambda = \frac{c}{\nu}$

Interaction of Energetic Electrons with Matter

When electrons hit matter:

- **Collision transfer** (99%): Energy transferred to other electrons → heat.
- **Radiative transfer** (1%): a) Inner-shell ionization → characteristic X-rays b) Braking near nucleus → bremsstrahlung radiation

Interaction of X-rays with Matter

Photoelectric effect Photon ejects an inner-shell electron:

$$E_e = h\nu - E_B$$

- Filling the vacancy emits characteristic X-rays.
- Alternatively produces Auger electrons.

Compton scattering Photon interacts with outer-shell electrons, losing energy and changing direction.

Generation of X-rays

X-rays are generated using an X-ray tube:

- Heated cathode emits electrons
- High voltage accelerates electrons
- Electrons hit anode → X-rays produced

Attenuation of Electromagnetic Radiation

Consider a narrow monoenergetic X-ray beam.

Let:
- $N(x)$ = number of photons
- $\mu(x)$ = linear attenuation coefficient

Photon loss:

$$dN = -\mu(x)Ndx$$

Divide and integrate:

$$d\frac{N}{N} = -\mu(x)dx$$

$\ln\left(\frac{N}{N_0}\right) = -\int \mu(x)dx$

Resulting intensity:

$N = N_0 \exp(-\int \mu(x)dx)$

Narrow Beam vs Broad Beam

- Broad beam: scattering contributes to detector signal.
- Monoenergetic assumption fails due to energy loss.

Solution:
- Collimation
- Narrow-beam geometry

Then attenuation law holds approximately.

Linear Attenuation Coefficient

$\mu(x)$ depends on:
- material
- photon energy

Higher $\mu \longrightarrow$ stronger attenuation.

Projection Radiography

Basic imaging equation:

$I = \int S(E) \exp(-\int \mu(x, E)dx)dE$

Assuming effective monoenergetic spectrum:

$I = I_0 \exp(-\int \mu(x)dx)$

Taking logarithm:

$-\ln\left(\frac{I}{I_0}\right) = \int \mu(x)dx$

Blurring in Projection Imaging

Sources of blur:
- Finite focal spot (penumbra)
- Detector blur
- Compton scattering outside field of view

Noise in Projection Imaging

Photon detection is a counting process:

$N \sim (N)$

Variance:

$\mathrm{Var}(N) = N$

Signal-to-noise ratio:

$\mathrm{SNR} = \frac{N}{\sqrt{N}} = \sqrt{N}$

To increase SNR:
- Increase photon count
- Use contrast agents

Tomography

Tomography = imaging by sectioning a volume.

From Greek:
- **tomos** = slice
- **grapho** = to write

Computed Tomography (CT)

Basic principle:
- Acquire many projections
- Different orientations around object
- Reconstruct cross-sectional image

CT Generations

- 1st generation: translate–rotate, pencil beam
- 2nd generation: fan beam, detector array
- 3rd generation: rotating source and detectors
- 4th generation: stationary detector ring

Image Formation in CT

Using attenuation model:

$I = I_0 \exp(-\int \mu(x) dx)$

Define projection value:

$p = -\ln\left(\frac{I}{I_0}\right) = \int \mu(x) dx$

Thus each projection is a line integral of μ.

Parallel-Ray Geometry

Parameterization:

$x(s) = s\cos(\theta) - t\sin(\theta)\ y(s) = s\sin(\theta) + t\cos(\theta)$

Projection:

$g(t, \theta) = \int \mu(x(s), y(s)) ds$

This is the **Radon transform**.

Sinogram

- $g(t, \theta)$ plotted over t and θ
- Each object point traces a sinusoid
- Sinogram contains all projection data

Backprojection

Idea:
- Smear each projection back over the image

Backprojection operator:

$f_{\mathrm{BP}(x,y)} = \int g(x\cos(\theta) + y\sin(\theta), \theta) d\theta$

Produces blurred image.

Fourier Slice Theorem

1D Fourier transform of projection:

$G(\omega, \theta) = F_1[g(t, \theta)]$

Equals slice of 2D Fourier transform of image:

[

$$F_2(\mu)(u, v)$$

]

with:

$u = \omega\cos(\theta)$ $v = \omega\sin(\theta)$

Filtered Backprojection (FBP)

Steps:
1. Filter projections with high-pass filter
2. Backproject filtered projections

Reconstruction: $\mu(x, y) = \int (g * h)(x\cos(\theta) + y\sin(\theta), \theta) d\theta$

where h is the reconstruction filter.

Iterative Reconstruction

- Start with initial guess
- Forward project
- Compare to measured data
- Update estimate

(Details skipped in lecture)

CT Artifacts

- **Aliasing**: insufficient number of projections
- **Beam hardening**: low-energy photons absorbed more strongly

Results in streaks and cupping artifacts.

Hounsfield Units

To standardize CT values:

[

$$HU = 1000\frac{\mu - \mu_{\text{water}}}{\mu_{\text{water}} - \mu_{\text{air}}}$$

]

Reference values:

- Air: −1000
- Water: 0
- Fat: −120 to −90
- Muscle: +35 to +55
- Bone: +300 to +1900
- White matter: +20 to +30
- Grey matter: +37 to +45

Summary

- CT reconstructs attenuation coefficients from projections
- Based on Radon transform and Fourier theory
- Filtered backprojection is classical reconstruction method
- Regularization and learning-based methods improve reconstruction

Learned Reconstruction

MRI

Image Registration

Image Segmentation

Federated Learning

Microscopy

Here's some example text. Notice how the section heading uses elegant spaced small caps.

A Subsection

Subsections use italic text for a subtle hierarchy.

Definition 8 (Important Concept) . A definition block with a distinctive left border. Use this to define key terms in your work.

Theorem 1 (Main Result) . A theorem block for stating important results. The numbering is automatic.

Example 1 — Practical Application . An example block with a subtle gray background. Use this to illustrate concepts with concrete examples.

Remark. A remark block for additional observations or notes that don't fit the formal structure of theorems and definitions.

Inline code looks like `this` , and code blocks are formatted cleanly:

```
def hello_world():
    """A simple function."""
    print("Hello, ClassicThesis!")
```

Tables and Figures

Item	Description	Value
Alpha	First item	100
Beta	Second item	200
Gamma	Third item	300

Table 1 : A sample table with clean styling.