Regularization Techniques for Tomography Problems

Chapter 12.1 and 12.2

Yiqiu Dong

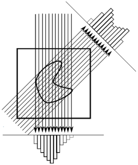
DTU Compute
Technical University of Denmark

16 Jan. 2023

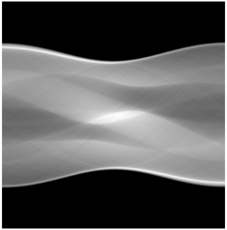
CT reconstruction

Object

$\bar{\boldsymbol{x}}$



$\Longrightarrow$



Measurements

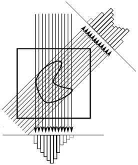
$\boldsymbol{b} = \mathcal{N}(\boldsymbol{A}\bar{\boldsymbol{x}})$

- **Forward Problem**: Send X-rays through the object at different angles, and measure the damping of X-rays.

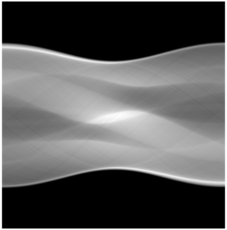
CT reconstruction

Object

$\bar{\boldsymbol{x}}$



$\Longleftarrow$

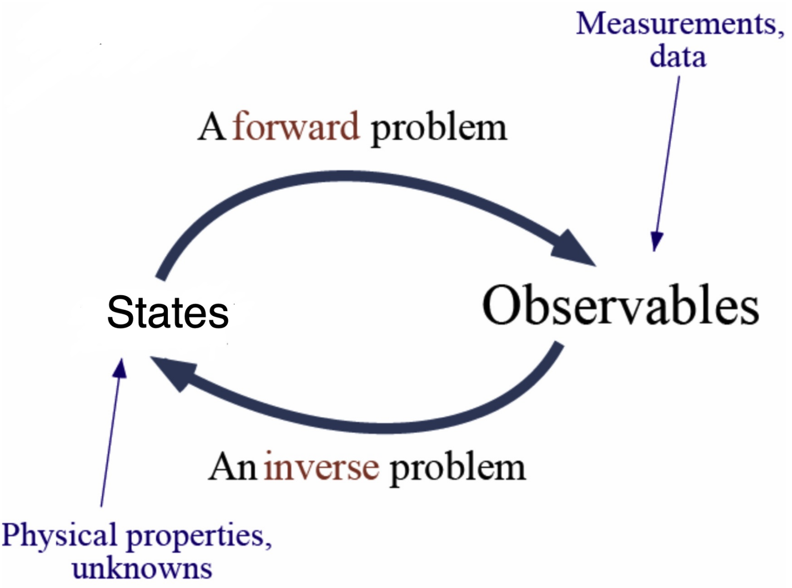


Measurements

$\boldsymbol{b} = \mathcal{N}(\boldsymbol{A}\bar{\boldsymbol{x}})$

- **Our Problem:** Reconstruct $\bar{\boldsymbol{x}}$ from $\boldsymbol{b}$ with given $\boldsymbol{A}$.
- It is a highly **ill-posed** inverse problem.

Inverse problems



Questions need be considered

- Why are inverse problems difficult?
  - Forward models are not explicitly invertible
  - Errors in the measurements (and also in the forward model) can lead to errors in the solution

Questions need be considered

- Why are inverse problems difficult?
  - Forward models are not explicitly invertible
    - ⋆ Existence: Does any state fit the measurement?
    - ⋆ Uniqueness: Is there a unique state vector fits the measurement?
  - Errors in the measurements (and also in the forward model) can lead to errors in the solution
    - ⋆ Stability: Can small changes in the measurement produce large changes in the solution?

Hadamard condition

A problem is called **well-posed** if

1. there exists a solution to the problem (existence),
2. there is at most one solution to the problem (uniqueness),
3. the solution depends continuously on the measurement (stability).

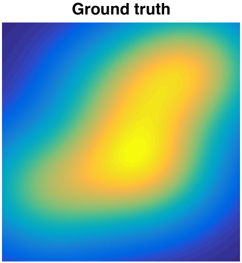
Otherwise the problem is called **ill-posed**.

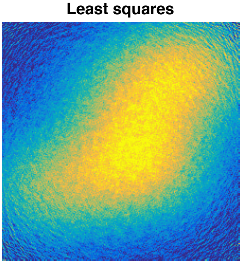
Example: Ill-posedness

- If too many measurements and no consistence, the solution of $\boldsymbol{Ax} = \boldsymbol{b}$ does not exist.
- If no enough measurements, the solution of $\boldsymbol{Ax} = \boldsymbol{b}$ is not unique.

Example: Ill-posedness

- If too many measurements and no consistence, the solution of $\boldsymbol{Ax} = \boldsymbol{b}$ does not exist.
- If no enough measurements, the solution of $\boldsymbol{Ax} = \boldsymbol{b}$ is not unique.
- Even we have a unique least-squares solution, it can be not good enough due to lack of the stability.





More questions need be considered

- Why are inverse problems difficult?

  $\Longleftarrow$ It's often ILL-POSED!
- How can we solve an ill-posed inverse problem?

More questions need be considered

- Why are inverse problems difficult?

  $\Longleftarrow$ It's often ILL-POSED!
- How can we solve an ill-posed inverse problem?
  - Does the measurements actually contain the information we want?
  - Which solution do we want?
  - The measurement may not be enough by itself to completely determine the unknown. What other prior information of the "unknown" do we have?

More questions need be considered

- Why are inverse problems difficult?

  $\Longleftarrow$ It's often ILL-POSED!

- How can we solve an ill-posed inverse problem?
  - Does the measurements actually contain the information we want?
  - Which solution do we want?
  - The measurement may not be enough by itself to completely determine the unknown. What other prior information of the "unknown" do we have?

  $\Longleftarrow$ We can use REGULARIZATION techniques!

Regularization techniques

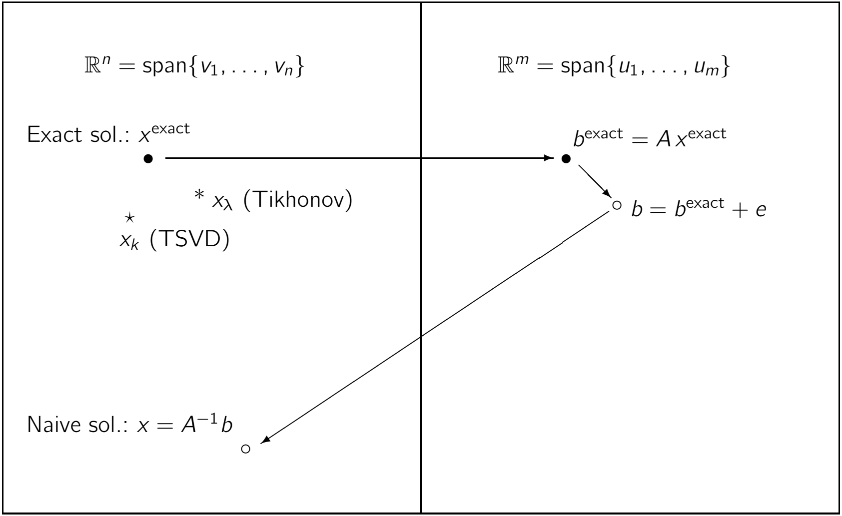
Consider to solve an ill-posed inverse problem:

$$\boldsymbol{b} = \mathcal{N}(\boldsymbol{A}\bar{\boldsymbol{x}})$$

Regularization: Approximate the inverse operator, $\boldsymbol{A}^{-1}$, by a family of stable operators $\mathcal{R}_\alpha$, where α is the regularization parameter.

We need: With the noise-free measurement we can find appropriate parameters α such that $\boldsymbol{x}_\alpha = \mathcal{R}_\alpha(\boldsymbol{b})$ is a good approximation of the true solution $\bar{\boldsymbol{x}}$.

Illustration of the need for regularization



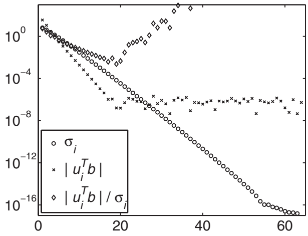
Truncated SVD

Considering the linear inverse problem

$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} \qquad \text{with } \boldsymbol{b} = \boldsymbol{A}\bar{\boldsymbol{x}} + \boldsymbol{e}.$$

Based on the SVD of $\boldsymbol{A}$, the "naive" solution is given by

$$\boldsymbol{x} = \boldsymbol{A}^{-1}\boldsymbol{b} = \sum_{i=1}^{I} \frac{\boldsymbol{u}_i^\top \boldsymbol{b}}{\sigma_i}\boldsymbol{v}_i = \bar{\boldsymbol{x}} + \sum_{i=1}^{I} \frac{\boldsymbol{u}_i^\top \boldsymbol{e}}{\sigma_i}\boldsymbol{v}_i$$



Truncated SVD

The solution of Truncated SVD is

$$\boldsymbol{x}_{\mathrm{TSVD}} = V\Sigma_k^{\dagger}U^{\top}\boldsymbol{b} = \sum_{i=1}^{k}\frac{\boldsymbol{u}_i^{\top}\boldsymbol{b}}{\sigma_i}\boldsymbol{v}_i$$

with $\Sigma_k^{\dagger} = \mathrm{diag}(\sigma_1^{-1}, \cdots, \sigma_k^{-1}, 0, \cdots, 0)$.

Truncated SVD

The solution of Truncated SVD is

$$\boldsymbol{x}_{\text{TSVD}} = V\Sigma_k^{\dagger}U^{\top}\boldsymbol{b} = \sum_{i=1}^{k} \frac{\boldsymbol{u}_i^{\top}\boldsymbol{b}}{\sigma_i}\boldsymbol{v}_i$$

with $\Sigma_k^{\dagger} = \mathrm{diag}(\sigma_1^{-1}, \cdots, \sigma_k^{-1}, 0, \cdots, 0)$.

- **Regularization parameter:**
  k, i.e, the number of SVD components.
- **Advantages:**
  - Intuitive
  - Easy to compute, *if we have the SVD*
- **Drawback:**
  - For large-scale problem, it is infeasible to compute the SVD

Tikhonov regularization

Idea: If we control the norm of the solution, then we should be able to suppress most of the large noise components.

Tikhonov regularization

Idea: If we control the norm of the solution, then we should be able to suppress most of the large noise components.

The Tikhonov solution $\boldsymbol{x}_{\mathrm{Tik}}$ is defined as the solution to

$$\min_{\boldsymbol{x}} \frac{1}{2}\|\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}\|_2^2 + \alpha\frac{1}{2}\|\boldsymbol{x}\|_2^2$$

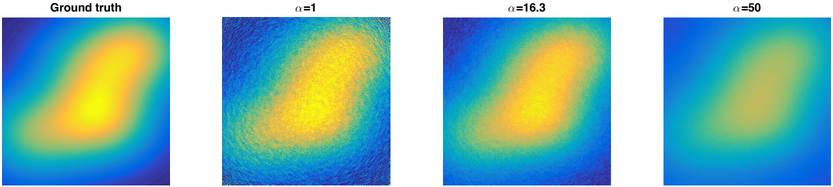
Tikhonov regularization

Idea: If we control the norm of the solution, then we should be able to suppress most of the large noise components.

The Tikhonov solution $\boldsymbol{x}_{\mathrm{Tik}}$ is defined as the solution to

$$\min_{\boldsymbol{x}} \frac{1}{2}\|\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}\|_2^2 + \alpha \frac{1}{2}\|\boldsymbol{x}\|_2^2$$

- **Regularization parameter:** α
- α large: strong regularity, over smoothing.
- α small: good fitting



Exercises

12.2 **Tikhonov Solution**

In this exercise, we study the property of the optimization problem:

$$\min_{\boldsymbol{x}} \frac{1}{2}\|\boldsymbol{Ax} - \boldsymbol{b}\|_2^2 + \alpha\frac{1}{2}\|\boldsymbol{x}\|_2^2$$

by calculating the gradient and the Hessian of the objective function.

The solution of Tikhonov regularization

Reformulate as a linear least squares problem

$$\min_{\boldsymbol{x}} \frac{1}{2} \left\| \begin{pmatrix} \boldsymbol{A} \\ \sqrt{\alpha}\boldsymbol{I} \end{pmatrix} \boldsymbol{x} - \begin{pmatrix} \boldsymbol{b} \\ 0 \end{pmatrix} \right\|_2^2$$

The solution of Tikhonov regularization

Reformulate as a linear least squares problem

$$\min_{\boldsymbol{x}} \frac{1}{2} \left\| \begin{pmatrix} \boldsymbol{A} \\ \sqrt{\alpha} \boldsymbol{I} \end{pmatrix} \boldsymbol{x} - \begin{pmatrix} \boldsymbol{b} \\ 0 \end{pmatrix} \right\|_2^2$$

The normal equation is

$$(\boldsymbol{A}^T \boldsymbol{A} + \alpha \boldsymbol{I})\, \boldsymbol{x} = \boldsymbol{A}^T \boldsymbol{b},$$

The solution of Tikhonov regularization

Reformulate as a linear least squares problem

$$\min_{\boldsymbol{x}} \frac{1}{2} \left\| \begin{pmatrix} \boldsymbol{A} \\ \sqrt{\alpha}\boldsymbol{I} \end{pmatrix} \boldsymbol{x} - \begin{pmatrix} \boldsymbol{b} \\ 0 \end{pmatrix} \right\|_2^2$$

The normal equation is

$$(\boldsymbol{A}^T\boldsymbol{A} + \alpha\boldsymbol{I})\,\boldsymbol{x} = \boldsymbol{A}^T\boldsymbol{b},$$

The solution is

$$\boldsymbol{x}_{\text{Tik}} = (\boldsymbol{A}^T\boldsymbol{A} + \alpha\boldsymbol{I})^{-1}\boldsymbol{A}^T\boldsymbol{b}$$

The solution of Tikhonov regularization

Reformulate as a linear least squares problem

$$\min_{\boldsymbol{x}} \frac{1}{2} \left\| \begin{pmatrix} \boldsymbol{A} \\ \sqrt{\alpha}\boldsymbol{I} \end{pmatrix} \boldsymbol{x} - \begin{pmatrix} \boldsymbol{b} \\ 0 \end{pmatrix} \right\|_2^2$$

The normal equation is

$$(\boldsymbol{A}^T\boldsymbol{A} + \alpha\boldsymbol{I})\,\boldsymbol{x} = \boldsymbol{A}^T\boldsymbol{b},$$

The solution is

$$\begin{aligned} \boldsymbol{x}_{\text{Tik}} &= (\boldsymbol{A}^T\boldsymbol{A} + \alpha\boldsymbol{I})^{-1}\boldsymbol{A}^T\boldsymbol{b} \\ &= V(\Sigma^2 + \alpha\boldsymbol{I})^{-1}\Sigma^\top U^\top \boldsymbol{b} \\ &= \sum_{i=1}^{n} \frac{\sigma_i(\boldsymbol{u}_i^\top \boldsymbol{b})}{\sigma_i^2 + \alpha}\boldsymbol{v}_i \end{aligned}$$

Compare with TSVD

- The solution of TSVD is

$$\boldsymbol{x}_{\text{TSVD}} = \sum_{i=1}^{k} \frac{\boldsymbol{u}_i^\top \boldsymbol{b}}{\sigma_i} \boldsymbol{v}_i$$

- The solution of Tikhonov regularization is

$$\boldsymbol{x}_{\text{Tik}} = \sum_{i=1}^{n} \frac{\sigma_i(\boldsymbol{u}_i^\top \boldsymbol{b})}{\sigma_i^2 + \alpha} \boldsymbol{v}_i$$

Compare with TSVD

- The solution of TSVD is

$$\boldsymbol{x}_{\text{TSVD}} = \sum_{i=1}^{k} \frac{\boldsymbol{u}_i^\top \boldsymbol{b}}{\sigma_i} \boldsymbol{v}_i = \sum_{i=1}^{n} \varphi_i^{\text{TSVD}} \frac{\boldsymbol{u}_i^\top \boldsymbol{b}}{\sigma_i} \boldsymbol{v}_i$$

with $\varphi_i^{\text{TSVD}} = \begin{cases} 1, & 1 \le i \le k, \\ 0, & k < i \le n. \end{cases}$

- The solution of Tikhonov regularization is

$$\boldsymbol{x}_{\text{Tik}} = \sum_{i=1}^{n} \frac{\sigma_i(\boldsymbol{u}_i^\top \boldsymbol{b})}{\sigma_i^2 + \alpha} \boldsymbol{v}_i = \sum_{i=1}^{n} \varphi_i^{\text{Tik}} \frac{\boldsymbol{u}_i^\top \boldsymbol{b}}{\sigma_i} \boldsymbol{v}_i$$

with $\varphi_i^{\text{Tik}} = \frac{\sigma_i^2}{\sigma_i^2+\alpha} \approx \begin{cases} 1, & \sigma_i \gg \sqrt{\alpha}\ , \\ \frac{\sigma_i^2}{\alpha}, & \sigma_i \ll \sqrt{\alpha}\ . \end{cases}$

Non-negativity and box constraints

- **Non-negativity constrained Tikhonov problem:**

$$\min_{\boldsymbol{x}\geq 0} \frac{1}{2}\|\boldsymbol{A}\boldsymbol{x}-\boldsymbol{b}\|_2^2+\alpha\frac{1}{2}\|\boldsymbol{x}\|_2^2$$

- **Box constrained Tikhonov problem:**

$$\min_{\boldsymbol{x}\in[a,b]^n} \frac{1}{2}\|\boldsymbol{A}\boldsymbol{x}-\boldsymbol{b}\|_2^2+\alpha\frac{1}{2}\|\boldsymbol{x}\|_2^2$$

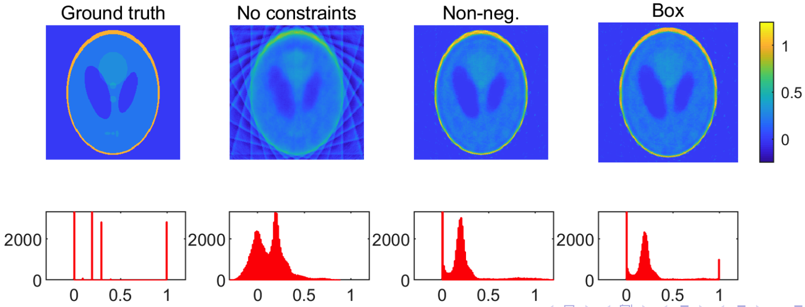
Non-negativity and box constraints

- **Non-negativity constrained Tikhonov problem:**

$$\min_{\boldsymbol{x} \geq 0} \frac{1}{2}\|\boldsymbol{A}\,\boldsymbol{x} - \boldsymbol{b}\|_2^2 + \alpha\,\frac{1}{2}\|\boldsymbol{x}\|_2^2$$

- **Box constrained Tikhonov problem:**

$$\min_{\boldsymbol{x} \in [a,b]^n} \frac{1}{2}\|\boldsymbol{A}\,\boldsymbol{x} - \boldsymbol{b}\|_2^2 + \alpha\,\frac{1}{2}\|\boldsymbol{x}\|_2^2$$



Exercises

12.3 **Influence of Regularization Parameters on Tikhonov Solutions**

We use a very small problem to study the influence of the regularization parameter.

Gaussian noise

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

where $\boldsymbol{e}$ denotes additive white Gaussian noise with zero mean and the covariance $\eta^2 \boldsymbol{I}_m$.

- All elements in $\boldsymbol{e}$ are independent.
- $\boldsymbol{e}$ is independent on $\bar{\boldsymbol{x}}$.
- Each element $\boldsymbol{e}_i$ can be seen as a Gaussian random variable with mean 0 and variance η^2.

Maximum likelihood estimate

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

where $\boldsymbol{e}$ denotes additive white Gaussian noise with zero mean and the covariance $\eta^2 \boldsymbol{I}_m$.

- The probability density for observing $\boldsymbol{b}$ given $\boldsymbol{x}$ is

$$\pi(\boldsymbol{b} \,|\, \boldsymbol{x}) = \pi(\boldsymbol{b} - \boldsymbol{A}\,\boldsymbol{x}) = \frac{1}{(\sqrt{2\pi}\eta)^m} \exp\left(-\frac{\|\boldsymbol{b} - \boldsymbol{A}\,\boldsymbol{x}\|_2^2}{2\eta^2}\right), \quad (1)$$

  which is called the *likelihood* of $\boldsymbol{x}$.

- **Maximum likelihood (ML) estimate** can be obtained by solving:

$$\max_{\boldsymbol{x}} \pi(\boldsymbol{b} \,|\, \boldsymbol{x}) \quad \Longleftrightarrow \quad \min_{\boldsymbol{x}} -\log\left(\pi(\boldsymbol{b} \,|\, \boldsymbol{x})\right).$$

- With the likelihood of $\boldsymbol{x}$ given in (1), we obtain the ML estimation problem

$$\min_{\boldsymbol{x}} \frac{1}{2}\|\boldsymbol{b} - \boldsymbol{A}\,\boldsymbol{x}\|_2^2 \,.$$

Maximum likelihood estimate

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

where $\boldsymbol{e}$ denotes additive white Gaussian noise with zero mean and the covariance $\eta^2 \boldsymbol{I}_m$.

- The probability density for observing $\boldsymbol{b}$ given $\boldsymbol{x}$ is

$$\pi(\boldsymbol{b} \,|\, \boldsymbol{x}) = \pi(\boldsymbol{b} - \boldsymbol{A}\,\boldsymbol{x}) = \frac{1}{(\sqrt{2\pi}\eta)^m} \exp\left(-\frac{\|\boldsymbol{b} - \boldsymbol{A}\,\boldsymbol{x}\|_2^2}{2\eta^2}\right), \quad (1)$$

  which is called the *likelihood* of $\boldsymbol{x}$.

- **Maximum likelihood (ML) estimate** can be obtained by solving:

$$\max_{\boldsymbol{x}} \pi(\boldsymbol{b} \,|\, \boldsymbol{x}) \quad \Longleftrightarrow \quad \min_{\boldsymbol{x}} -\log\left(\pi(\boldsymbol{b} \,|\, \boldsymbol{x})\right).$$

- With the likelihood of $\boldsymbol{x}$ given in (1), we obtain the ML estimation problem

$$\min_{\boldsymbol{x}} \frac{1}{2}\|\boldsymbol{b} - \boldsymbol{A}\,\boldsymbol{x}\|_2^2 \,.$$

Maximum likelihood estimate

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

where $\boldsymbol{e}$ denotes additive white Gaussian noise with zero mean and the covariance $\eta^2 \boldsymbol{I}_m$.

- The probability density for observing $\boldsymbol{b}$ given $\boldsymbol{x}$ is

$$\pi(\boldsymbol{b} \,|\, \boldsymbol{x}) = \pi(\boldsymbol{b} - \boldsymbol{A}\,\boldsymbol{x}) = \frac{1}{(\sqrt{2\pi}\eta)^m} \exp\left(-\frac{\|\boldsymbol{b} - \boldsymbol{A}\,\boldsymbol{x}\|_2^2}{2\eta^2} \right), \qquad (1)$$

  which is called the *likelihood* of $\boldsymbol{x}$.
- **Maximum likelihood (ML) estimate** can be obtained by solving:

$$\max_{\boldsymbol{x}} \pi(\boldsymbol{b} \,|\, \boldsymbol{x}) \quad \Longleftrightarrow \quad \min_{\boldsymbol{x}} -\log\left(\pi(\boldsymbol{b} \,|\, \boldsymbol{x})\right).$$

- With the likelihood of $\boldsymbol{x}$ given in (1), we obtain the ML estimation problem

$$\min_{\boldsymbol{x}} \frac{1}{2} \|\boldsymbol{b} - \boldsymbol{A}\,\boldsymbol{x}\|_2^2 \ .$$

MAP esitmate

To obtain a stable solution, we can incorporate prior information on $\bar{\boldsymbol{x}}$ by applying Bayes formula:

$$\pi(\boldsymbol{x} \mid \boldsymbol{b}) = \frac{\pi(\boldsymbol{b} \mid \boldsymbol{x})\, \pi_{\text{prior}}(\boldsymbol{x})}{\pi(\boldsymbol{b})}\ .$$

- $\pi(\boldsymbol{x} \mid \boldsymbol{b})$ is the posterior.
- $\pi(\boldsymbol{b} \mid \boldsymbol{x})$ is the likelihood.
- $\pi_{\text{prior}}(\boldsymbol{x})$ is the prior probability density of $\boldsymbol{x}$.
- $\pi(\boldsymbol{b})$ is the prior probability density of $\boldsymbol{b}$.

MAP esitmate

To obtain a stable solution, we can incorporate prior information on $\bar{\boldsymbol{x}}$ by applying Bayes formula:

$$\pi(\boldsymbol{x} \mid \boldsymbol{b}) = \frac{\pi(\boldsymbol{b} \mid \boldsymbol{x})\, \pi_{\text{prior}}(\boldsymbol{x})}{\pi(\boldsymbol{b})} \,.$$

- $\pi(\boldsymbol{x} \mid \boldsymbol{b})$ is the posterior.
- $\pi(\boldsymbol{b} \mid \boldsymbol{x})$ is the likelihood.
- $\pi_{\text{prior}}(\boldsymbol{x})$ is the prior probability density of $\boldsymbol{x}$.
- $\pi(\boldsymbol{b})$ is the prior probability density of $\boldsymbol{b}$.

Maximum a posteriori (MAP) estimate can be obtained by solving:

$$\begin{aligned} \max_{\boldsymbol{x}} \pi(\boldsymbol{x} \mid \boldsymbol{b}) &\iff \max_{\boldsymbol{x}} \frac{\pi(\boldsymbol{b} \mid \boldsymbol{x})\, \pi_{\text{prior}}(\boldsymbol{x})}{\pi(\boldsymbol{b})}, \\ &\iff \min_{\boldsymbol{x}} -\log(\pi(\boldsymbol{b} \mid \boldsymbol{x})) - \log(\pi_{\text{prior}}(\boldsymbol{x})), \end{aligned}$$

Example

If we have

- the likelihood: $\pi(\boldsymbol{b}\,|\,\boldsymbol{x}) = \frac{1}{(\sqrt{2\pi}\eta)^m}\exp\left(-\frac{\|\boldsymbol{A}\,\boldsymbol{x}-\boldsymbol{b}\|_2^2}{2\eta^2}\right)$ and
- the prior: $\pi_{\text{prior}}(\boldsymbol{x}) = \frac{1}{(\sqrt{2\pi}\beta)^n}\exp(-\frac{1}{2\beta^2}\|\boldsymbol{x}\|_2^2)$ (Gaussian distribution),

then the **MAP estimate** can be obtained by solving

$$\min_{\boldsymbol{x}} \frac{1}{2}\|\boldsymbol{b}-\boldsymbol{A}\,\boldsymbol{x}\|_2^2 + \alpha\,\frac{1}{2}\|\boldsymbol{x}\|_2^2$$

with $\alpha = \eta^2/\beta^2$.

Example

If we have

- the likelihood: $\pi(\boldsymbol{b} \,|\, \boldsymbol{x}) = \frac{1}{(\sqrt{2\pi}\eta)^m} \exp\left(-\frac{\|\boldsymbol{A}\,\boldsymbol{x} - \boldsymbol{b}\|_2^2}{2\eta^2}\right)$ and
- the prior: $\pi_{\text{prior}}(\boldsymbol{x}) = \exp(-\frac{1}{\beta} J(\boldsymbol{x}))$ (Gibbs prior) with $\beta > 0$,

then the **MAP estimate** can be obtained by solving

$$\min_{\boldsymbol{x}} \frac{1}{2} \|\boldsymbol{b} - \boldsymbol{A}\,\boldsymbol{x}\|_2^2 + \alpha\, J(\boldsymbol{x})$$

with $\alpha = \eta^2/\beta$.

Example

If we have

- the likelihood: $\pi(\boldsymbol{b} \,|\, \boldsymbol{x}) = \frac{1}{(\sqrt{2\pi}\eta)^m} \exp\left(-\frac{\|\boldsymbol{A}\,\boldsymbol{x} - \boldsymbol{b}\|_2^2}{2\eta^2}\right)$ and
- the prior: $\pi_{\text{prior}}(\boldsymbol{x}) = \exp(-\frac{1}{\beta} J(\boldsymbol{x}))$ (Gibbs prior) with $\beta > 0$,

then the **MAP estimate** can be obtained by solving

$$\min_{\boldsymbol{x}} \frac{1}{2}\|\boldsymbol{b} - \boldsymbol{A}\,\boldsymbol{x}\|_2^2 + \alpha\, J(\boldsymbol{x})$$

with $\alpha = \eta^2/\beta$.

- The term $\frac{1}{2}\|\boldsymbol{b} - \boldsymbol{A}\,\boldsymbol{x}\|_2^2$ is called the *data-fidelity* term.
- The term $J(\boldsymbol{x})$ is called the *regularization* term.
- $\alpha > 0$ is the regularization parameter.

Poisson Measurements in X-ray

The measured transmission I_i in a single detector element follows a Poisson distribution $\mathcal{P}(I_0 \exp(-\boldsymbol{r}_i^T \boldsymbol{x}))$:

$$\pi(I_i \mid \boldsymbol{x}) = \frac{\left(I_0 \exp(-\boldsymbol{r}_i^T \boldsymbol{x})\right)^{I_i}}{I_i!} \exp\left(-I_0 \exp(-\boldsymbol{r}_i^T \boldsymbol{x})\right),$$

where $\boldsymbol{r}_i^T$ with $i = 1, \cdots, m$ denotes the row of the system matrix $\boldsymbol{A}$.

Poisson Measurements in X-ray

The measured transmission I_i in a single detector element follows a Poisson distribution $\mathcal{P}(I_0 \exp(-\boldsymbol{r}_i^T \boldsymbol{x}))$:

$$\pi(I_i \,|\, \boldsymbol{x}) = \frac{\left(I_0 \exp(-\boldsymbol{r}_i^T \boldsymbol{x})\right)^{I_i}}{I_i!} \exp\left(-I_0 \exp(-\boldsymbol{r}_i^T \boldsymbol{x})\right),$$

where $\boldsymbol{r}_i^T$ with $i = 1, \cdots, m$ denotes the row of the system matrix $\boldsymbol{A}$.

- The likelihood: $\pi(\boldsymbol{I} \,|\, \boldsymbol{x}) = \prod_{i=1}^{m} \pi(I_i \,|\, \boldsymbol{x})$.
- The ML estimate ($\boldsymbol{b} = -\log(\boldsymbol{I}/I_0)$):

$$\arg\min_{\boldsymbol{x}} \; -\log\left(\pi(\boldsymbol{b} \,|\, \boldsymbol{x})\right) \iff \arg\min_{\boldsymbol{x}} \; \exp(-\boldsymbol{b})^T \boldsymbol{A}\,\boldsymbol{x} + 1^T \exp(-\boldsymbol{A}\,\boldsymbol{x}).$$

- The MAP estimate: $\arg\min_{\boldsymbol{x}} \exp(-\boldsymbol{b})^T \boldsymbol{A}\,\boldsymbol{x} + 1^T \exp(-\boldsymbol{A}\,\boldsymbol{x}) + \alpha J(\boldsymbol{x})$.

Exercise

- **12.1 Quadratic Approximation for Poisson Noise.**
  Use the second-order Taylor expansion of

$$D_i(\tau) = \exp(-b_i)\,\tau + \exp(-\tau), \qquad i = 1, \ldots, m\ ,$$

  to verify that the ML estimation problem can be approximated by the weighted quadratic problem

$$\min_{\boldsymbol{x}} \frac{1}{2}(\boldsymbol{A}\,\boldsymbol{x} - \boldsymbol{b})^T W\,(\boldsymbol{A}\,\boldsymbol{x} - \boldsymbol{b})$$

  with $W = \mathrm{diag}(\exp(-\boldsymbol{b}))$.