Image reconstruction

Part II
Inverse problems - A methodology

Céline Theys

Laboratoire Lagrange
UNS/CNRS/OCA

Équipe Signal & Image

Besançon 2014

Outline

A simple example: optical images deconvolution

Mathematical Model

$$y(n, m) = \sum_{k,l} h(k, l)x(n - k, m - l) + (b(n, m)) \qquad (1)$$

The impulse response h includes:

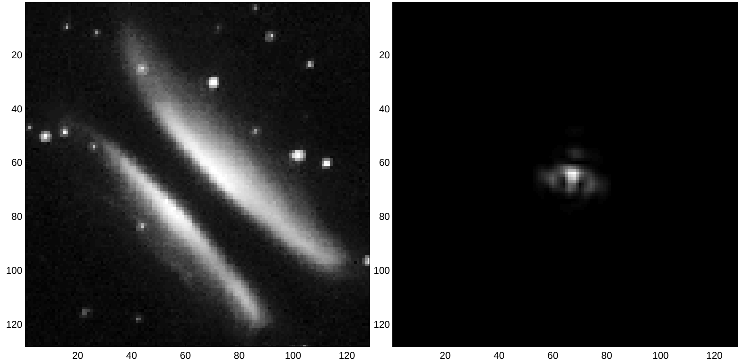
- The limited aperture of the optical system.
- A move of the object during the exposition.
- Atmospheric turbulences.
- All of these phenomena (astrophysical: 1 + 3).

b: measure and model errors

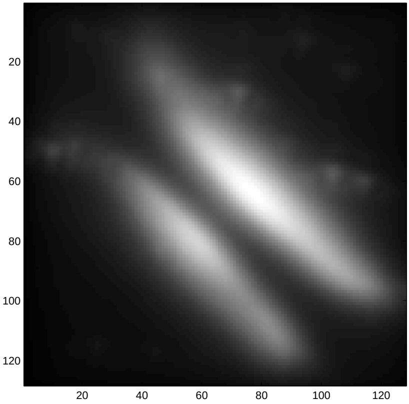
x: object

y: image

HST data, Psf


20
40
60
80
100
120
20 40 60 80 100 120
20
40
60
80
100
120
20 40 60 80 100 120

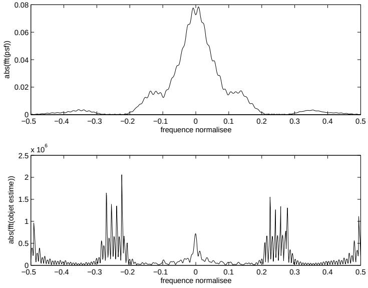
Result of the convolution



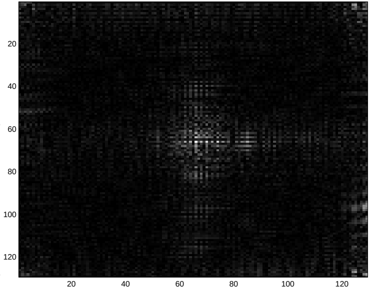
Solution in the Fourier plan

$$Y(f_x, f_y) = H(f_x, f_y)X(f_x, f_y) \quad (2)$$

$|H(f_x = 0, f_y)|$, $|X(f_x = 0, f_y)|$



Estimated Object



Remarks

$$X(f_x, f_y) = \frac{Y(f_x, f_y)}{H(f_x, f_y)} - \left(\frac{B(f_x, f_y)}{H(f_x, f_y)}\right) \quad (3)$$

- Lowpass system with transmission zeros.
- Without noise, amplification of the components at frequencies corresponding to the transmission zeros.
- With noise, amplification of the noise at the same frequencies.
- *Signal processing: if the noise is Gaussian, it is "the best" solution (unbiased and minimum variance estimator)*

Conclusion: ill posed problem!!

Ill posed problems

$$y = S[x] \qquad (4)$$

y: measures; x: object; $S[]$: known transformation.
Problem: recover x from y?

Well posed problems: Hadamard conditions

1. Existency. The inverse operator exists, there is at least one solution x.
2. Unicity. The solution x is unique.
3. Stability. The solution is stable.

First kind Fredholm equation

Monodimensional case ($k(s,r)$: kernel, $\mathbf{x}(\mathbf{r})$: unknown)

$$y(s) = \int_a^b k(s,r)x(r)dr \tag{5}$$

For translation invariant kernel, **convolution equation**:

$$y(s) = \int_{-\infty}^{\infty} k(s-r)x(r)dr \tag{6}$$

First kind Fredholm equation

Bidimensional case (optical imagery)

$$y(s,t) = \int_a^b \int_c^d k(s,t,r,v)x(r,v)drdv \tag{7}$$

For translation invariant kernel, **bi-dimensional convolution equation**:

$$y(s,t) = \int_a^b \int_c^d k(s-r,t-v)x(r,v)drdv \tag{8}$$

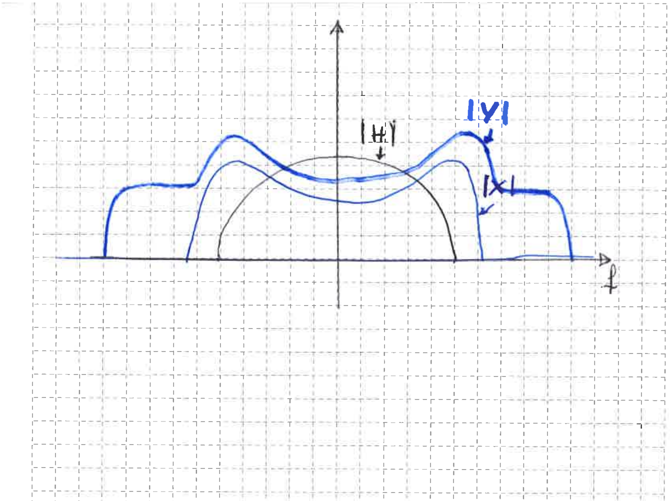
And separable:

$$y(s,t) = \int_a^b \int_c^d l(s-r)m(t-v)x(r,v)drdv \tag{9}$$

Existency

Existency

"If an unlimited bandwidth error (white noise for example) or outside the bandwidth is added on y then there is no solution



Unicity

Riemann-Lebesgue theorem

$$\lim_{\alpha \to \infty} \int_a^b k(s,r)(x(r) + \sin(\alpha r))dr = \int_a^b k(s,r)x(r)dr \qquad (10)$$

"If high component frequencies are added to x, they are removed by the transformation, i.e x is recovered up to high frequencies"

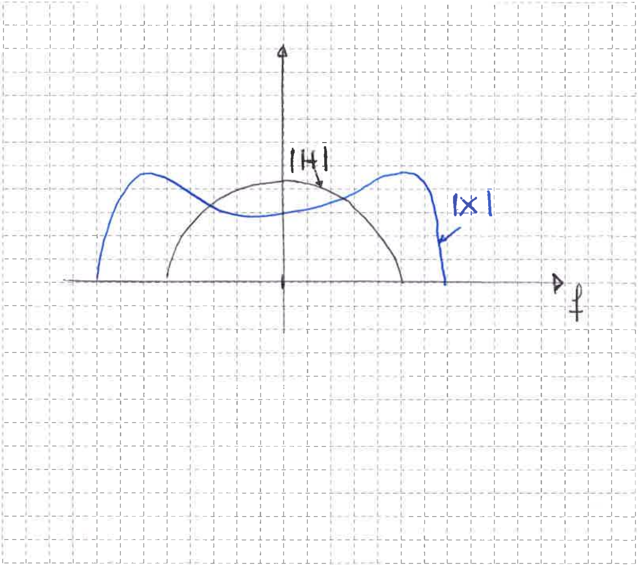
More generally $x(r) = u(r) + v(r)$

$$\int_a^b k(s,r)(u(r) + v(r))dr = \int_a^b k(s,r)u(r)dr \qquad (11)$$

with $v(r)$ orthogonal to $k(r,s)$.

"x is recovered up to component frequencies outside the bandwidth of the instrument"

Unicity



Discretization

x and y belong to finite dimensional spaces and k is then a matrix H:

$$\mathbf{y} = H\mathbf{x} + (\mathbf{b}) \tag{12}$$

Structure of H - One dimensional convolution

$$y_n = \sum_{i=-Q}^{P} x_{n-i} h_i \tag{13}$$

One value of $y \rightarrow P + Q + 1$ values of x
Shift of one for each new value of y
N values of $y \rightarrow P + Q + N$ values of x

Structure of H - One dimensional convolution

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{pmatrix} = \begin{pmatrix} h_P & \dots & h_0 & \dots & h_{-Q} & 0 & \dots \\ 0 & h_P & \dots & h_0 & \dots & h_{-Q} & 0 \\ \vdots & \ddots & \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & h_P & \dots & h_0 & \dots & h_{-Q} \end{pmatrix} \begin{pmatrix} x_{-P+1} \\ \vdots \\ x_0 \\ x_1 \\ \vdots \\ x_{1+Q} \\ \vdots \\ x_N \\ \vdots \\ x_{N+Q} \end{pmatrix}$$

Using the DFT

- Underdetermined system:
  $dim(\mathbf{x}) = N + P + Q > dim(\mathbf{y}) = N$.
- Add $P + Q$ rows to H to get a square circulant matrix
  $dim(H) = (N + P + Q) \times (N + P + Q)$.
- Then we can compute $\mathbf{y}$ by
  1. DFT of $\mathbf{x}$ (dimension L).
  2. DFT of the first row of H, (dimension L).
  3. Product term by term of the two DFTs.
  4. IDFT of the result.
  5. Extraction of the N "correct" values from the $N + P + Q$ values of $\mathbf{y}$.

Structure of H - Two dimensional convolution

$$y_{n,k} = \sum_{i=-Q_1}^{P_1} \sum_{j=-Q_2}^{P_2} h_{i,j} x_{(n-i,k-j)} \tag{14}$$

$$y_{1,1} = \sum_i \left(h_{(i,-Q_2)} x_{(1-i,1+Q_2)} + h_{(i,-Q_2+1)} x_{(1-i,Q_2)} + \ldots + h_{(i,P_2)} x_{(1-i,1-P_2)} \right) \tag{15}$$

One value of $y_{n,k} \rightarrow (P_1 + Q_1 + 1)(P_2 + Q_2 + 1)$ values of $x_{n,k}$
N^2 values $\rightarrow (P_1 + Q_1 + N^2)(P_2 + Q_2 + N^2)$ values of $x_{n,k}$

$$
\begin{pmatrix} y_{1,1} \\ y_{1,2} \\ \vdots \\ y_{N,N} \end{pmatrix} =
\begin{pmatrix}
h_{P_1,P_2} & \dots & h_{P_1,-Q_2} & \dots & h_{-Q_1,-Q_2} & 0 & \dots & \\
0 & h_{P_1,P_2} & \dots & h_{P_1,-Q_2} & \dots & h_{-Q_1,-Q_2} & \dots & \\
 & \ddots & & \ddots & & \ddots & & \ddots \\
0 & \dots & h_{P_1,P_2} & \dots & h_{P_1,-Q_2} & \dots & h_{-Q_1,-Q_2} &
\end{pmatrix}
\begin{pmatrix} x_{(1-P_1,1-P_2)} \\ \vdots \\ x_{(1+Q_1,1+Q_2)} \\ x_{(2-P_1,2-P_2)} \\ \vdots \\ x_{(2+Q_1,2+Q_2)} \\ \vdots \\ x_{(N+Q_1,N+Q_2)} \end{pmatrix} \quad (16)
$$

Solutions

Whatever dimensions and without noise or unknown noise

$$\mathbf{x}_{est} = arg \min_{\mathbf{x}} ||\mathbf{y} - H\mathbf{x}||^2 = (H^T H)^{-1} H^T \mathbf{y} \qquad (17)$$

With i.i.d noise

$$\mathbf{x}_{est} = (H^T H)^{-1} H^T \mathbf{y} \qquad (18)$$

With known autocorrelation matrix

$$\mathbf{x}_{est} = (H^T R^{-1} H)^{-1} H^T R^{-1} \mathbf{y} \qquad R = E[\mathbf{b}\mathbf{b}^T] \qquad (19)$$

$(H^T H)^{-1} H^T = H^\dagger$: generalized inverse matrix

Conclusion: bad conditioning

In the discretised case, there is always a solution, the problem is the inversion of H or $H^T H \rightarrow H$ **is bad conditioned**.

Condition number

$$c = \frac{\lambda_{max}}{\lambda_{min}} \tag{20}$$

λ_{max}: maximal eigenvalue, λ_{min}: minimal eigenvalue
well-conditioned: $c \approx 1$, bad-conditioned: $c \gg 1$,
(introducing example: $c = 4.10^5$)

Solutions to bad conditioning

Truncated eigendecomposition

- The eigendecomposition allows to increase the conditioning by suppressing the smaller eigenvalues.
- If the λ_i are arranged in descending order and the eigenvectors are arranged in the order of the eigenvalues, one can simply truncate the decomposition by stopping the summation to main eigenvalues
- $(H^T H)^{-1} = \sum_{i=1}^{\mathbf{K}} \frac{1}{\lambda_i^2} \mathbf{v}_i \mathbf{v}_i^T$

Problems

- Choice of the truncation?
- No physical information on the object ...

Adding a priori - Regularization

- By adding a priori, we give up the true solution, anyway the true solution is lost.
- We search a "stabilized" solution.
- Signal processing: biased estimator.

Minimization of a composite criterion

$$J(\mathbf{x}) = J_1(\mathbf{y}, \mathbf{x}) + \gamma J_2(\mathbf{x}) \tag{21}$$

J_1: fidelity term, J_2 **a priori term on x or regularization term**, γ: regularization coefficient.
Constraint: $J(\mathbf{x})$ must be convex!

Regularization

Smoothness constraint to reduce the amplification of the noise.

Tikhonov regularization

$$J_2(\mathbf{x}) = ||D\mathbf{x}||^2 \tag{22}$$

With D a linear operator.

- $J_2(\mathbf{x}) = ||\mathbf{x}||^2 = ||\mathbf{x} - 0||^2$ "we search to minimize the norm of the solution"

Regularization

- $J_2(\mathbf{x}) = ||C\mathbf{x}||^2$ with C first derivative operator "we search to minimize the norm of the derivative of the solution"

$$C = \begin{pmatrix} 1 & -1 & 0 & \dots & \dots \\ 0 & 1 & -1 & 0 & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots \end{pmatrix}$$

- $J_2(\mathbf{x}) = ||L\mathbf{x}||^2$ with L second derivative operator "we search to minimize the norm of the second derivative of the solution"

$$L = \begin{pmatrix} 1 & -2 & 1 & 0 & \dots \\ 0 & 1 & -2 & 1 & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots \end{pmatrix}$$

Regularization

$$J_2(\mathbf{x}) = ||\mathbf{x} - \mathbf{x}_{priori}||^2{}_W = (\mathbf{x} - \mathbf{x}_{priori})^T W^{-1} (\mathbf{x} - \mathbf{x}_{priori}) \quad (23)$$

- This distance corresponds to a Gaussian a priori on $\mathbf{x}$ of mean $\mathbf{x}_{priori}$ and covariance matrix W.
- If the fidelity term is also quadratic, then:

$$\begin{aligned} J(\mathbf{x}) &= ||\mathbf{y} - H\mathbf{x}||^2{}_R + \gamma ||\mathbf{x} - \mathbf{x}_{priori}||^2{}_W \\ &= (\mathbf{y} - H\mathbf{x})^T R^{-1} (\mathbf{y} - H\mathbf{x}) + \gamma (\mathbf{x} - \mathbf{x}_{priori})^T W^{-1} (\mathbf{x} - \mathbf{x}_{priori}) \end{aligned}$$

Regularization

Explicit minimizer

$$\mathbf{x}_{est} = (H^T R^{-1} H + \gamma W^{-1})^{-1} (H^T R^{-1} \mathbf{y} + \gamma W^{-1} \mathbf{x}_{priori}) \quad (24)$$

For $\gamma = 0$, we find the generalized inverse solution
For $\gamma \to \infty$, $\hat{\mathbf{x}} \to \mathbf{x}_{priori}$

Remark: We add γ on the diagonal of $H^T H$ (if W is diagonal)

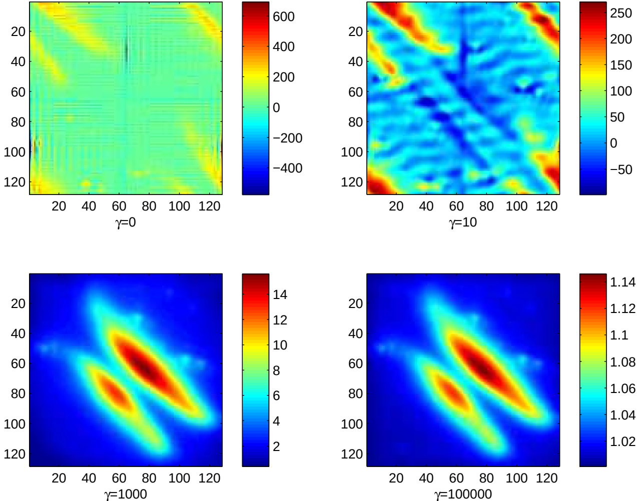
Simulation example

```
>>image=conv2(psf128,objet19,'same'); >>imagebruitee=image+randn(128); >>gamma=10

>>dirac=zeros(128); >>dirac(64,64)=1;

>>objetest=ifft2((fft2(imagebruitee)+gamma*fft2(ones(128)))

./(fft2(psf128)+gamma*abs(fft2(dirac)));
```



Regularization

How to choose γ?

- Find a value of the regularization coefficient such that:

$$\mathbf{x}_{est} = arg\min_{\mathbf{x}}(J_1(\mathbf{y}, \mathbf{x}) + \gamma J_2(\mathbf{x})) \tag{25}$$

be as close as possible to $\mathbf{x}$:

$$\gamma_{est} = arg\min_{\gamma}(||\mathbf{x}_{est} - \mathbf{x}||^2) \tag{26}$$

- No explicit solution in practice since $\mathbf{x}$ is not known.
- There are some results in the quadratic case: "curve in L", cross validation ...

Iterative methods?

- No explicit solution
- Explicit solution but high order inversion
- Inequalities constraints (positivity ...) $\rightarrow$ non linear equation

$$[\nabla J(x^\star)]x^\star = 0 \tag{27}$$

*: optimum

- Stop before convergence! (without regularization)

Descent algorithm

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \alpha^{(k)}\mathbf{d}^{(k)} \tag{28}$$

$\mathbf{d}^{(k)}$: descent direction/ $J(\hat{\mathbf{x}}^{(k+1)}) \leq J(\hat{\mathbf{x}}^{(k)})$
$\alpha^{(k)}$: step size

Richardson Lucy algorithm

- Poisson distribution:

$$\mathbf{y} = \mathcal{P}(H\mathbf{x}) \tag{29}$$

- Fidelity term:

$$J(\mathbf{x}) = \sum_{i=1}^{N} ((H\mathbf{x})_i - y_i \log(H\mathbf{x})_i) \tag{30}$$

- Descent direction:

$$\mathbf{d}^{(k)} = -[\nabla_{\mathbf{x}} J(\mathbf{x})].\mathbf{x} = H^T(\mathbf{y}./H\mathbf{x} - 1_N).\mathbf{x} \quad \textit{and} \quad \alpha^{(k)} = \alpha = 1 \tag{31}$$

Richardson Lucy algorithm

$$x_i^{(k+1)} = x_i^{(k)} [H^T]_i \frac{y_i}{[H\mathbf{x}^{(k)}]_i} \tag{32}$$

Image Space Reconstruction algorithm

- Additive Gaussian noise:

$$\mathbf{y} = H\mathbf{x} + \mathbf{b} \qquad y_i = (H\mathbf{x})_i + b_i \qquad R = E(\mathbf{b}^T\mathbf{b}) \tag{33}$$

- Fidelity term:

$$J(\mathbf{x}) = ||\mathbf{y} - H\mathbf{x}||_R^2 = (\mathbf{y} - H\mathbf{x})^T R^{-1} (\mathbf{y} - H\mathbf{x}) \tag{34}$$

- Descent direction:

$$\mathbf{d}^{(k)} = -\frac{[\nabla_{\mathbf{x}} J(\mathbf{x})]}{H^T R H\mathbf{x}}.\mathbf{x} = \left(\frac{-H^T R H\mathbf{x} + H^T R\mathbf{y}}{H^T R H\mathbf{x}} \right).\mathbf{x} \tag{35}$$

For $\alpha^{(k)} = \alpha = 1$, ISRA algorithm

$$x_i^{(k+1)} = x_i^{(k)} \frac{[H^T R\mathbf{y}]_i}{[H^T R H\mathbf{x}^{(k)}]_i} \tag{36}$$

Conclusion

Observation by an optical system is an ill-posed problem

→ Direct inversion gives irrelevant result
→ **Regularization** = introduction of smoothness a priori
→ **Iterative methods** for explicit gradient, non linear constraints or/and high dimensional problem
→ **If no invariant psf, inversion of a very high dimensional matrix**